PROCEEDINGS
OF THE
NATIONAL ACADEMY OF SCIENCES
Volume 31 September 15, 1945 Number 9
Copyright 1945 by the National Academy of Sciences

*PURE STRAIN MICE BORN TO HYBRID MOTHERS FOLLOWING OVARIAN TRANSPLANTATION**

BY W. L. RUSSELL AND JANE GOODRICH HURST†

ROSCOE B. JACKSON MEMORIAL LABORATORY, BAR HARBOR, MAINE

Communicated July 20, 1945

It is well established that many characters in mammals are affected by naturally occurring variation in the uterus or, to use a more inclusive term, the prenatal maternal environment. Although this factor is regarded as environmental as far as the embryo is concerned, it is itself presumably subject to both environmental and genetic factors (Russell[1]).

The work of Wright,[2] Wright and Chase,[3] Green [4] and others,[5] who made statistical analyses of variation within inbred strains of mammals, clearly shows that the characters which they studied, polydactyly, white spotting, number of vertebrae, etc., were affected by environmentally determined variation in the prenatal maternal environment.

The effect of genetically determined variation in the prenatal maternal environment on character variation has, however, received little attention, perhaps because of a lack of easy methods of evaluating it. Evidence that this factor may be important is provided by the increasing number of cases in which a difference has been observed between reciprocal F_1 hybrids obtained by crossing inbred strains. Russell and Green[6] have reported one such case for number of lumbar vertebrae in mice. Similar data from other crosses have been obtained and are in preparation for publication. The mere discovery of such a difference does not, however, furnish proof of a difference between the maternal environments of the parental strains: without further analysis the difference between the hybrids could equally well be attributed to the egg cytoplasm or, in the heterogametic sex, to sex linked genes.

Two methods have been developed which can be used to test directly for differences between prenatal maternal environments. The first has so far been utilized for this purpose on a large scale only by Fekete and Little.[7] In their work, fertilized ova are removed from the oviduct of one inbred

strain of mice and transferred, by means of a fine glass pipette, to the uterus of another inbred strain. The resulting offspring thus develop from the chromosomes and cytoplasm of the first strain and in the uterine environment of the second.

The second method, developed recently by Robertson,[8] makes ingenious use of transplantation of ovaries within an inbred strain maintained by forced heterozygosis. In Robertson's experiments, ovaries from yellow (A^yA^w) mice were transplanted to agouti (A^wA^w) mice of the same inbred strain. By mating these, and unoperated yellow females, with yellow males he was able to compare the development of the lethal homozygous A^yA^y type in the uterus of an agouti with its development in the uterus of a yellow.

The two methods differ in that the transplantation of ovaries brings the foster mother's environment into play at an earlier stage. Whether this is important or not is not known, but a combination of the two methods on the same material could be used to investigate this point. The methods differ also in their practical limitations. The small average number of offspring obtained from operations with transferred ova would make the collection of the large number of individuals often required for the study of quantitative variation somewhat tedious. Ovarian transplantation, while more promising in this respect, seemed to be limited to studies of single factor genetic differences within inbred strains maintained by forced heterozygosis. It was presumably not available for the investigation of multiple factor differences between strains because of the failure of tissue grafts in foreign hosts. That interstrain transplants of mouse ovaries are unsuccessful has been shown by Robertson[8] in a total of 36 operations and by Mr. Ralph Kellogg† at this laboratory in a total of 23 operations.

An attempt to find a method by which ovarian transplantation could be used to measure the effect of multiple genetic differences in the maternal environment led to the experiments described in this paper. It occurred to us that if ovaries could be transplanted from an inbred strain to the F_1 hybrid of that and another strain, a type of transplant that is usually successful, this would provide a method for comparing pure strain and hybrid maternal environments.

For the first experiment, strain dba ovaries were transplanted to dba/C3H F_1 hybrids which were then mated with dba males. These strains were chosen because the dba strain (genotype *ddbbaa*) differs by three recessive color factors from the C3H strain (*DDBBAA*), a fact which, as shown in figure 1, would help in distinguishing successful operations from those cases in which the hybrid host ovary regenerated. Thus offspring from the transplanted ovary would all be dilute brown non-agouti (*ddbbaa*), while only one-eighth of the offspring from a host's regenerated ovary would be of this color. In all operations, except those on females Nos. 162

and 163 (see table 1), we followed the essentials of the technique described by Robertson.[8] The left ovary, capsule and oviduct of each host were removed. Then, 7 to 15 days later, a second operation was performed in which the host's right ovary was removed from its capsule and the donor ovary implanted in its place. In female No. 162 the first operation was omitted, one host ovary being left intact, and in No. 163 there was no interval between the two operations. The latter variation in technique is discussed later in this paper.

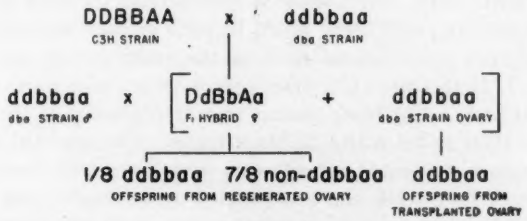


FIGURE 1

Of the 22 operated hybrid females, 9 had no offspring and one produced a single litter which was eaten shortly after birth. The remaining 12 gave the results shown in table 1. Females Nos. 17, 23 and 58 clearly show re-

TABLE 1

NUMBER AND TYPE OF OFFSPRING FROM MATINGS OF dba MALES WITH dba/C3H F_1 HYBRID FEMALES INTO WHICH dba OVARIES HAD BEEN IMPLANTED

Serial No. of hybrid host	17	23	25	26	45	55	58	65	162	163	193	197
ddbbaa offspring	..	..	2	2	19	4	..	18	1	12	9	8
Non-*ddbbaa* offspring	6	3	8	17	..	24	7	4	11	..	...	...

generation of the host ovary and failure of the transplanted ovary. Females Nos. 25, 26, 55 and 162 also indicate failure of the graft, for the proportion of *ddbbaa* offspring produced by each of these females is not significantly different from one-eighth. In females Nos. 45, 163, 193 and 197 the transplanted ovary was clearly functional and there was apparently no regeneration of the host ovary. The remaining female, No. 65, produced 4 non-*ddbbaa* offspring which must have come from a regenerated host ovary, but the high proportion of *ddbbaa* offspring, 18 out of 22, indicates that the transplanted ovary was also functioning.

Robertson[8] does not mention the possibility of regeneration occurring along with a functional graft in his results, but his ratio of 38 yellow to 62 agouti, where a 1:1 ratio was expected from the transplanted ovary, indicates that regeneration probably occurred in some of his females and a ratio of 38 yellow to 52 agouti from those females which produced at least one yellow offspring, though not significantly different from a 1:1 ratio,

does differ in the direction that favors the possibility of regeneration having occurred even in the presence of a functional graft.

In our work, in any case, subsequent experiments have shown that this type of result is not uncommon and apparently cannot be avoided even by careful operative technique. This is unfortunate because in the above experiments, for example, even where the success of the transplant was established by the fact that only *ddbbaa* animals were born, one cannot be certain that any particular *ddbbaa* offspring was not produced from a regenerated host ovary. With hybrids from strains differing by a smaller number of marker genes there would be even greater uncertainty, as we found in a series of operations made by the junior author and Miss Barbara Perry.† In this series C57 black (genotype *aa*) mice were crossed with C3H agouti (*AA*). C57 black ovaries were transplanted to the F_1 hybrids which were then mated with C57 black males. The expected ratio in the offspring from a regnerated hybrid ovary is thus one half agouti (*Aa*) and one half black (*aa*), while the transplanted ovary would yield only black offspring. Of the 23 animals that had young, only one gave reasonably clear-cut results. This produced 3 agouti and 18 black offspring, indicating a combination of successful transplant and regenerated host ovary. Of the remaining 22, 8 gave only black offspring (a total of 26) and 14 gave both agouti and black (totaling 56 and 70, respectively). The total of 56 agouti to 96 black produced by these 22 females indicates that the ovarian grafts were functional in at least some, but the number of offspring was not large enough to be sure of this in any particular case.

It is apparent that the ever-present possibility of regeneration of the host ovary occurring along with a successful graft is a serious limitation. The difficulty, of course, resides in the fact that the hybrid carries the marker genes of both parental strains and will segregate each parental combination in at least some of its germ cells. To get around this difficulty a method was worked out which makes use of an inbred strain of mice which we have called "stock 129" and which was originally obtained from Prof. L. C. Dunn. This strain was inbred with forced heterozygosis to carry both chinchilla (c^{ch}) and albino (c^a) alleles. It also carries pink-eye and light-bellied agouti genes, but this can be disregarded in the present discussion. It is now maintained, as shown in figure 2 (*a*), by brother-sister matings of the heterozygotes ($c^{ch}c^a \times c^{ch}c^a$), thus producing, in addition to the heterozygotes, the two homozygous combinations $c^{ch}c^{ch}$ and c^ac^a. All three types are phenotypically distinguishable. Two ways in which this strain is being used in ovarian transplant experiments are shown in figure 2. In figure 2 (*b*) the F_1 hybrid is obtained by mating any full colored, *CC*, strain with the stock 129 $c^{ch}c^{ch}$ type. Ovaries from stock 129 c^ac^a animals are transplanted to the Cc^{ch} hybrids which are then mated with stock 129 $c^{ch}c^{ch}$ males. Offspring from a regnerated host ovary will thus be either Cc^{ch} or $c^{ch}c^{ch}$, genotypes

which are phenotypically distinguishable from the $c^{ch}c^{a}$ combination produced by a successful ovarian graft. The origin of every individual born to the hybrid can, therefore, be determined. Figure 2 (*c*) shows a similar method that can be applied to a cross with any albino strain. The system of matings with the strains shown in figure 2 can be worked out in other ways, but the two examples given serve to illustrate the principle.

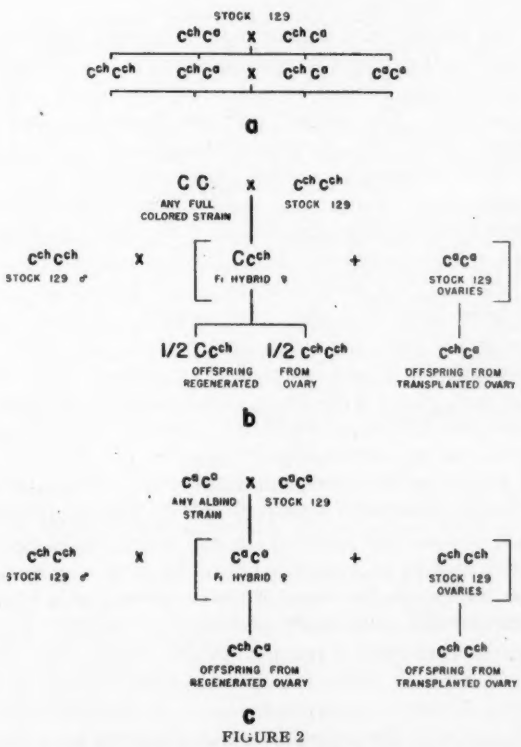


FIGURE 2

(*a*) System of maintaining stock 129. (*b*, *c*) Two ways in which this stock can be used in ovarian transplantations to hybrid females so that offspring of the grafted ovary can be distinguished from offspring of a regenerated host ovary.

There are other possibilities. Thus $c^{ch}c^{ch}$ or $c^{a}c^{a}$ strains could be used just as well as the *CC* and $c^{a}c^{a}$ strains, and strains carrying other genes maintained in forced heterozygosis could be used in place of stock 129. As a control for possible effects of the operation itself, transplants can be

made both ways between the two homozygous types of the 129 stock or whatever stock is used in its place.

In the above systems, pure strain mice are raised in hybrid hosts by taking the transplanted ovary, as well as the male with which its host is mated, from the same strain. It is, however, equally possible to use males from other strains, thus producing hybrid offspring raised in hybrid uteri. One case of this is reported below. Of course the choice of strains is again restricted to those carrying suitable marker genes.

Summarizing the above, it is clear that, within the limitation of having to choose appropriate strains of animals, a method is provided by which ovarian transplantation can be used for comparing pure strain and hybrid maternal environments. The method could also be used in other fields of biology wherever, for example, an experiment requires a means of distinguishing between the functioning of host and graft ovarian tissue.

In our experiments, ovaries from stock 129 have been transplanted to hybrids of this stock and four other stocks: C3H, dba, C57 black and Bagg albino. The first three are full colored (*CC*) strains and were used as shown in figure 2 (*b*), while the Bagg albino was mated as in figure 2 (*c*). Functional grafts have been obtained in all of these hybrids. In most cases the hybrids were mated to stock 129 males, thus yielding pure stock 129 animals raised in four different hybrid maternal environments. Some of the 129/C57 hybrids were mated with Bagg albino males to produce, from the transplanted ovary, 129/Bagg hybrids developed in 129/C57 uteri.

Details of technique, and findings in regard to the effect of various factors on the success of the transplantations, will be presented in another paper, but it seems desirable to mention here the following departure from Robertson's technique. He removed one host ovary 7 to 10 days before the transplantation in order to produce hypertrophy of the remaining ovary and of its capsule into which the ovary of the donor was to be placed, but he questions whether this is necessary and suggests testing it by removal of both host ovaries at the time of transplantation. The success obtained with female No. 163 (table 1), which was operated upon in this way, shows that hypertrophy of the site of implantation is not an essential factor. In view of this finding, most of the experiments with stock 129 have been made by removing both ovaries of a host and implanting donor ovaries in their place in the same operation. In our work, operations have been more successful with this than with Robertson's method, possibly because two graft ovaries per operated female are used instead of one.

At the present time about 40% of over 100 operations made by the senior author on stock 129 hybrids have resulted in successful grafts. Nearly one-half of these have shown at least some regeneration of host ovarian tissue. The number of offspring obtained from the grafted ovaries has been very high in some cases. The highest number recorded for a single female is 80.

This was a 129/Bagg hybrid containing stock 129 ovaries and mated with a stock 129 male. The number of young is higher than has ever been recorded at this laboratory for a straight stock 129 mating. (The operation was made by Miss Winona Hinkley† who assisted in the early work with the 129 stock.)

A large series of experiments is being conducted to compare the maternal environments of the 129 stock and some of its hybrids, using number of vertebrae as a character. It has already been stated that differences in number of lumbar vertebrae obtained in reciprocal hybrids of certain strains indicate that this character may be affected by the maternal environment.

Summary.—The only way, so far reported, of testing directly for the effect of multiple gene, or strain, differences in maternal environments is by the somewhat tedious process of transferring fertilized ova from one strain to another. The work reported here provides a method of using ovarian transplantation, for this purpose, at least as far as a comparison of pure strain and hybrid maternal environments is concerned.

Offspring were obtained from pure strain ovaries transplanted to hybrid females. Six different host-donor combinations, involving a total of three inbred strains and six hybrids, were tried and all were successful. In many cases a successful graft was accompanied by a regenerated hybrid host ovary. In early experiments it was not possible to distinguish offspring of the grafted and regenerated ovaries except on a statistical basis, but mating plans were devised using strains with suitable marker genes in such a way that it is possible to determine the origin of all the offspring by their phenotype.

* Aided by a grant from the Rockefeller Foundation and by grants to the Roscoe B. Jackson Memorial Laboratory from the National Cancer Institute, the Jane Coffin Childs Memorial Fund for Medical Research, the Anna Fuller Fund and the International Cancer Research Foundation.

† Summer student at the Roscoe B. Jackson Memorial Laboratory.

[1] Russell, W. L., Chap. 10 in Staff of the Roscoe B. Jackson Memorial Laboratory *Biology of the Laboratory Mouse*, Blakiston Co., Philadelphia, 497 pp. (1941).

[2] Wright, S., *Genetics*, **19**, 506–536 (1934).

[3] Wright, S., and Chase, H. B., *Ibid.*, **21**, 758–787 (1936).

[4] Green, E. L., *Ibid.*, **26**, 192–222 (1941).

[5] See Green[4] for references.

[6] Russell, W. L., and Green, E. L., *Genetics*, **28**, 87 (1943).

[7] Fekete, E., and Little, C. C., *Cancer Res.*, **2**, 525–530 (1942).

[8] Robertson, G. G., *Jour. Exp. Zool.*, **89**, 197–232 (1942).

EXPERIMENTS ON SEXUAL ISOLATION IN DROSOPHILA. V. THE EFFECT OF VARYING PROPORTIONS OF DROSOPHILA PSEUDOOBSCURA AND DROSOPHILA PERSIMILIS ON THE FREQUENCY OF INSEMINATION IN MIXED POPULATIONS

BY HOWARD LEVENE AND TH. DOBZHANSKY*

COLUMBIA UNIVERSITY

Communicated July 19, 1945

Introduction.—When females of *Drosophila pseudoobscura* and *Drosophila persimilis* are placed together with males of one of these species, it is usually found that a greater proportion of the conspecific than of the alien females are inseminated. The nature of the stimuli that lead to this result is still obscure, although Mayr and Dobzhansky[1] have shown that the degree of the preference varies depending upon the history of the individual flies involved and upon the environment. The experiments to be reported in the present article are concerned with the effects of the relative numbers of the conspecific and of the alien females available to the males on the frequencies of homogamic and heterogamic matings. Perhaps the simplest of the many hypotheses that may be constructed about the matings which take place in mixed populations is that when a female and a male meet, there is a certain definite probability that they will mate, depending on their respective species but not on what other males and females may be found in the same medium. On the other hand, it is also possible that the flies are stimulated by the proximity of conspecific individuals of the opposite sex in such a manner that they become more receptive to mating with alien as well as with conspecific partners than they would otherwise be. Still another possibility is that only homogamic matings take place while conspecific individuals of the opposite sex are available, and that heterogamic matings occur only, or predominantly, when the possibility of homogamic mating is excluded or remote. The above possibilities are not necessarily mutually exclusive, and various intermediate situations may be encountered. It is also fully conceivable that different pairs of species and races might behave quite differently in this respect.

Methods and Procedure.—The orange strain of *D. pseudoobscura* and the Stony Creek strain of *D. persimilis* were used in all experiments.[1] In the main body of seven experiments, 10 freshly hatched *D. persimilis* males were placed in each vial together with the following freshly hatched females: (1) 10 *persimilis*, (2) 10 *pseudoobscura*, (3) 10 *persimilis* and 10 *pseudoobscura*, (4) 20 *persimilis* and 10 *pseudoobscura*, (5) 40 *persimilis* and 10 *pseudoobscura*, (6) 10 *persimilis* and 20 *pseudoobscura*, and (7) 10 *persimilis* and 40 *pseudoobscura*. In other words, only *D. persimilis* males were used in all vials and their numbers were always ten per vial; the numbers and

the species of the females were varied. The vials were kept in an incubator at 25½°C. for approximately 4 days, whereupon the females were dissected and the presence or absence of sperm in their seminal receptacles was determined by microseopic examination. These experiments were done in September–October, 1944, with the assistance of Mr. George Streisinger, whose help we wish to acknowledge. The results are summarized as experiments 1 to 7 in table 1.

TABLE 1

NUMBERS OF INSEMINATED FEMALES OBSERVED WHEN 10 *D. persimilis* MALES WERE PLACED WITH n_c *D. persimilis* FEMALES AND n_a *D. pseudoobscura* FEMALES

EXPT. NUMBER	NO. OF CONSPECIFIC PER VIAL	NO. OF ALIEN PER VIAL	*D. persimilis* FERTILIZED	*D. persimilis* UNFERTILIZED	*D. persimilis* TOTAL	*D. pseudoobscura* FERTILIZED	*D. pseudoobscura* UNFERTILIZED	*D. pseudoobscura* TOTAL	ISOLATION INDEX	ISOLATION RATIO
x	n_{cx}	n_{ax}	F_{cx}	$N_{cx}-F_{cx}$	N_{cx}	F_{ax}	$N_{ax}-F_{ax}$	N_{ax}	b_x	g_x
1	10	0	....* 115* 80.41*	.... 28 19.59	143					
2	0	10				... 28 18.30	.... 125 81.70	153		
3	10	10	143.67 141 79.21	34.33 37 20.79	178	27.34 39 22.54	145.66 134 77.46	173	0.5569	0.2846
4	20	10	149.68 150 78.53	41.33 41 21.47	191	14.27 13 13.98	78.73 80 86.02	93	0.6978	0.1780
5	40	10	204.67 204 54.69	168.33 169 45.31	373	9.67 11 12.22	80.33 79 87.78	90	0.6347	0.2235
6	10	20	57.03 60 58.82	44.97 42 41.18	102	22.99 17 8.10	187.01 193 91.90	210	0.7581	0.1376
7	10	40	87.72 90 69.77	41.28 39 30.23	129	70.17 64 12.14	456.83 463 87.86	527	0.7035	0.1741
8	0	10				... 49 37.98	.... 80 62.02	129		
9	10	10	.... 122 95.31	.... 6 4.69	128	... 67 51.54	.... 63 48.46	130		

* For each experiment, the top line gives the "expected number" of flies on the basis of H2′ (see text), the middle line gives the observed numbers, and the bottom line gives the observed percentages.

Additional data were secured in April, 1945, with the aid of a slightly different method. Males and females of *D. persimilis* and *D. pseudoobscura* were aged in the absence of individuals of the opposite sex for a week to ten days, whereupon 10 *persimilis* males were placed in vials overnight with 10 *pseudoobscura* females or with 10 *pseudoobscura* and 10 *persimilis* females. (Experiments 8 and 9 in table 1.)

Similar experiments were performed on two geographic strains of *D. prosaltans*, but these will be discussed in a separate section.

When freshly hatched females and males were kept together for 3–4 days before dissection, the results obtained in different vials of the same experiment frequently showed statistically significant heterogeneity, while the experiments with the aged flies did not. The heterogeneity is, of course, a complicating factor in the analysis of the data. There are several probable sources of this heterogeneity. Since dissection and examination of the sperm receptacles is a rather laborious operation, females and males in some vials were kept together somewhat longer than in others. Some flies came from culture bottles in the early and others in late stages of hatching. Since the experiments of a given series extended for about two months, some variation in the food and other environmental factors may have occurred.

Analysis.—Throughout this discussion we will assume that there are 10 males of a given species or strain present in a standard vial, along with n_c conspecific[2] females and n_a alien females. Of these f_c and f_a, respectively, will be fertilized. Let N_c equal the sum of the n_c for a given experiment (i.e., N_c is the total number of conspecific females in all the vials of a given composition), and let N_a, F_c and F_a be the corresponding sums for n_a, f_c and f_a.

If π_c is the probability that any particular conspecific female will be inseminated during the course of the experiment, then π_c will be some function of n_c, n_a and of t, the time that the flies are left together. We denote this function by $\pi_c(n_c, n_a, t)$. We will include in t also such factors as temperature, age of the flies, food, etc., so that t represents what may be called the "physiological time." Similarly $\pi_a(n_c, n_a, t)$ is the probability that an alien female will be inseminated. Evidently π_c and π_a approach zero as the number of females per vial becomes large, since there must be some upper limit to the number of females a male can inseminate. Because of the variability from vial to vial of the per cent fertilized, as above discussed, the data do not give very precise information on the form of the functions π_c and π_a. However, from examination of the results for the individual vials it is clear that there is no very great variation in π_c over the range of values of n_c and n_a used, with the possible exception of a decrease in π_c in the experiment with 40 conspecific and 10 alien females. Apparently the decrease in percentage of fertilization does not become marked until considerably higher densities of females are reached than were used here. Since any analytic function is approximately linear for small changes in the variables, we may suppose that π_c depends on some linear combination of n_c and n_a when n_c and n_a do not vary too widely. We may represent this combination as $n' = dn_c + en_a$ where d and e are constants. If, now, $e/d = 0$, π_c depends only on the number of conspecific females,

while if $e/d = 1$ it depends only on the total number of females. Since π_c apparently decreases for $n_c = 40$, $n_a = 10$ and not for $n_c = 10$, $n_a = 40$, e/d seems to be a positive fraction less than one.

Stalker[3] has proposed an "isolation index" $\beta = (\pi_c - \pi_a)/(\pi_c + \pi_a)$. An equivalent but somewhat simpler index is $\gamma = \pi_a/\pi_c$. They may be interchanged by the formulas $\beta = (1 - \gamma)/(1 + \gamma)$ and $\gamma = (1 - \beta)/(1 + \beta)$. If only homogamic matings occur, $\beta = 1$ and $\gamma = 0$; if there is no discrimination, $\beta = 0$ and $\gamma = 1$; and if only heterogamic matings occur, $\beta = -1$ and $\gamma = \infty$. It is somewhat easier to interpret γ than β, as the probability of an alien female being inseminated is simply γ times the probability of a conspecific female being inseminated. We shall call γ the "isolation ratio."

We shall now consider the implications of the hypotheses, discussed above, about the manner in which the mating preferences become effective. If males do not mate with alien females unless conspecific females are unavailable, we should expect to find no inseminated aliens until nearly all the conspecific females have been inseminated. Since this is not so, we must look for some further explanation. It might be that only a small proportion of the females present are receptive at any one time. This seems rather farfetched; it might, however, be tested by seeing what happens when only one or two males are present. On the other hand, the hypothesis might be altered to give a hypothesis which we will call H1 by supposing that the probability π_a of an alien female being inseminated is not zero but has some low value when conspecific females are available and has a higher value when only aliens are available. If we also suppose that the presence of conspecific females excites males and makes them more likely to mate, we will have a situation where π_a has one low value in the absence of conspecific females, another low value when conspecific females are available, and a higher value when conspecific females have been available but are so no longer. The balance between these opposing tendencies would then determine whether π_a would be increased or decreased when the number of conspecific females n_c was increased from zero to, e.g., ten. Introducing a second subscript to refer to the experiment number, the question is whether π_{a3}/π_{a2} is greater than or less than one. The data give[4] $p_{a3}/p_{a2} = 1.23$, but this does not differ significantly from one. This comparison was repeated, using aged flies to decrease the heterogeneity as explained above. When these vials were tested for heterogeneity they gave $\chi^2 = 19.56$ with 28 degrees of freedom, for $P = 0.85$, so that these results are much more reliable. This time $p_{a9}/p_{a8} = 1.36$ and this is significant, with $\chi^2 = 4.81$ and $P = 0.03$.

Before discussing H1 further we will consider an alternative hypothesis, H2. We now suppose that the isolation ratio $\gamma = \pi_a(n_c, n_a, t)/\pi_c(n_c, n_a, t)$ is a constant independent of n_c, n_a, and t. Evidently this cannot hold for

all possible n_c, n_a and t; for example, if t is very large and n_c and n_a small, then π_c, π_a and hence γ will tend to 1; however γ may be essentially constant over a wide range of values. We shall not specify any definite forms for the functions π_c and π_a but merely suppose that in the xth experiment π_c has a fixed value π_{cx} and π_a has a fixed value π_{ax}. We temporarily ignore the heterogeneity of the data. Hypothesis H2 now reduces to H2′, where we assume that π_a/π_c has a common value γ_x for each vial of experiment x, and the hypothesis states that all the γ_x are equal. We represent the common value of the γ_x by γ.

Evidently γ_x involves a comparison between two kinds of females and hence cannot be computed for experiments 1 and 2, and therefore they cannot be used in testing H2′. For the other experiments, under the assumption only (i.e., when the γ_x need not be equal), the best estimates for π_{cx}, etc., are p_{cx}, etc., where p_{cx} is the proportion of conspecific females fertilized in all vials of the xth experiment; and the best estimate of γ_x is $g_x = p_{ax}/p_{cx}$. The g_x are given in table 1.

When H2′ is true, the p's are no longer the best estimates. We may now obtain π_{ax} from the relation that under H2′, $\pi_{ax} = \gamma\pi_{cx}$ but we still require joint estimates of γ, π_{c3}, π_{c4}, ..., π_{c7}. The method of maximum likelihood[5] leads to the set of six non-linear equations

$$\left.\begin{aligned} &\sum_{x=3}^{7}\left[\frac{F_{ax}}{\gamma} - \frac{\pi_{cx}(N_{ax} - F_{ax})}{1 - \gamma\pi_{cx}}\right] = 0 \\ &\left[\frac{F_{cx} + F_{ax}}{\pi_{cx}} - \frac{N_{cx} - F_{cx}}{1 - \pi_{cx}} - \frac{\gamma(N_{ax} - F_{ax})}{1 - \gamma\pi_{cx}}\right] = 0 \quad (x = 3, 4, \ldots, 7). \end{aligned}\right\} \quad (1)$$

It is not practicable to solve these equations exactly, but a method of successive approximations gives the estimate $g = 0.1958$ and estimates of the π_{cx} which lead to the "expected values" that are given in table 1. The ordinary χ^2 test is not strictly applicable here because the constraints are non-linear. However the likelihood ratio statistic,[6] λ, is easily calculated and, for samples as large as this, $-2 \log_e \lambda$ should be distributed approximately as χ^2 with 4 degrees of freedom.[7]

To compute $-2 \log_e \lambda$ we need the expression

$$(-2 \log_e 10) \sum_{x=3}^{7} [(F_{cx} + F_{ax}) \log \pi_{cx} + F_{ax} \log \gamma_x + (N_{cx} - F_{cx}) \log (1 - \pi_{cx}) + (N_{ax} - F_{ax}) \log (1 - \gamma_x\pi_{cx})]. \quad (2)$$

(All logarithms are to the base 10 unless otherwise indicated.) Now let L_1 be the value of (2) when we replace γ_x by γ and then replace γ, π_{c3}, π_{c4}, ..., π_{c7} by the numerical estimates obtained fom equations (1). Also, let L_2 be the value of (2) when we replace γ_3, ..., γ_7, π_{c3}, ..., π_{c7}, by the

sample values given in table 1. Then $-2 \log_e \lambda = L_1 - L_2$ We obtain the result $-2 \log_e \lambda = 9.0$, corresponding to a probability of about 0.061.

Discussion.—In nature, H2 would mean essentially that there is one probability of copulation when a male meets one of his own females and another (smaller) probability when he meets an alien female, while under H1 we have the added complication that his willingness to mate with the alien is reduced by the presence of receptive conspecific females. It should be noted that we have spoken throughout as if it were the male who did the choosing. This has been done solely for economy of words. Actually it is not known whether it is the males or the females or both which exercise the discrimination. It may be that the alien females reject the males, and that H1 means that the males are more likely to persevere if no conspecific females are present. The end result will be the same in any case. The important thing is that, under H2, π_a and π_c both decrease at the same rate, as n_c increases, while, under H1, π_a decreases more rapidly than π_c so that γ decreases. The data also require the assumption that males become excited if conspecific females are present, and then are more likely to mate with alien females. This additional effect could be superimposed on either H1 or H2, but would have to be more extreme to override the effect of H1. Since there is no other evidence in favor of H1, it seems that the simpler hypothesis H2 should be accepted. We have seen that, ignoring heterogeneity, the probability of the data arising by chance under H2′ is about 0.06. The heterogeneity would tend to increase the probability of more extreme values of the statistic $-2 \log_e \lambda$, and hence to further reduce the need of rejecting H2′. On the other hand, the heterogeneity would also tend to overshadow any real differences in the γ_x and hence make the test less sensitive. Table 1 shows on its face that the g_x do vary over a range of 2 to 1, but in an irregular manner rather than in the regular way which we would expect under H1. The actual variation in the γ_x could be even greater than this, but it is not likely that very large systematic changes occur. All things considered it seems best to accept the simpler hypothesis H2, since it is not contradicted by the data, and indeed seems to fit them better than H1. In any event the simple hypothesis that males only mate with alien females when their own females are unavailable seems untenable, and its modification, H1, is simply a more complicated version of H2.

Results with Two Strains of D. prosaltans.—Experiments were also carried out with the same number of flies as in experiments 1 to 7, but using flies of the Chilpancingo (Mexico) strain of *D. prosaltans* instead of *D. persimilis* and flies of the Bertioga (Brazil) strain of *D. prosaltans* instead of *D. pseudoobscura*. As shown by Dobzhansky and Streisinger,[8] males of both these strains exhibit a strong preference for mating with Chilpancingo females. Freshly hatched flies were kept together in an incubator at 25½°C. for approximately 3 days, whereupon all the females were dissected and their

seminal receptacles examined for sperm. These experiments were performed in August–September, 1944. The results are summarized as experiments 1′ to 7′ in table 2.

TABLE 2

D. prosaltans. NUMBERS OF INSEMINATED FEMALES OBSERVED WHEN 10 CHILPANCINGO MALES WERE PLACED WITH n_c CHILPANCINGO FEMALES AND n_a BERTIOGA FEMALES

EXPT. NUMBER	NO. OF CONSPECIFIC PER VIAL	NO. OF ALIEN PER VIAL	CHILPANCINGO			BERTIOGA			ISOLATION INDEX	ISOLATION RATIO
			FERTILIZED	UNFERTILIZED	TOTAL	FERTILIZED	UNFERTILIZED	TOTAL		
x	n_{cx}	n_{ax}	F_{cx}	$N_{cx}-F_{cx}$	N_{cx}	F_{ax}	$N_{ax}-F_{ax}$	N_{ax}	b_x	g_x
1′	10	0	57*	33	90					
			63.33*	36.67						
2′	0	10				14	86	100		
						14.00	86.00			
3′	10	10	72	27	99	2	95	97	0.945	0.0284
			72.73	27.27		2.06	97.94			
4′	20	10	116	58	174	1	86	87	0.966	0.0172
			66.67	33.33		1.15	98.85			
5′	40	10	156	240	396	1	97	98	0.950	0.0259
			39.39	60.61		1.02	98.98			
6′	10	20	57	22	79	3	152	155	0.948	0.0268
			72.15	27.85		1.94	98.06			
7′	10	40	64	32	96	12	381	393	0.912	0.0458
			66.67	33.33		3.05	96.95			

* For each experiment, the top line gives the observed number of flies and the bottom line gives the observed percentages.

In these experiments the observed frequencies of heterogamic fertilization are so low that their sampling fluctuations would overshadow any reasonable changes in the γ_x. All that can be said is that the data in experiments 3′ through 7′ do not contradict H2′. For such low γ values large numbers of vials would have to be used to obtain good results. The only significant result is the considerable increase of heterogamic matings when no conspecific females are present. We have $p_{a3}'/p_{a2}' = 0.147$, with $\chi^2 = 7.43$, using Yates' correction for continuity, and $P = 0.006$. This is probably a real effect, although it might be due to the heterogeneity between vials. If real, this effect is more consistent with H1 than with H2. It could be at least partly explained under H2 as follows. Suppose π_c and π_a depend on an "effective population density" with an added female affecting π_a in proportion to her chance of insemination; i.e., if, as suggested above, we consider π_a to depend on a linear combination, $n' = dn_c + en_a = d[n_c + (e/d)n_a]$, of n_c and n_a, then e/d will be equal to γ. Now if $\pi_a(n', t)$ decreases rapidly as n' increases from 0 to 2 or 4, and thereafter decreases less rapidly, then if γ is small, as it is here, 10 γ will be small, and increasing n_c from 0 to 10 will reduce π_a considerably. There is as yet no experimental evidence on the behavior of π_c and π_a for small n_c by which this conjecture could be tested.

Thus the difference in behavior in the two series of experiments need not necessarily imply different causal systems. However, it would not be surprising if they were different, since in one case we are dealing with sexual isolation between two species and in the other with a one-way sexual preference between two strains of a single, different species.

Summary.—Results obtained by placing *D. persimilis* males with varying proportions of *D. persimilis* and *D. pseudoobscura* females admit of the hypothesis that the ratio of the probability of a heterogamic mating to the probability of a homogamic mating is a fixed constant independent of these proportions. However, the possibility of some decrease in this ratio when many *D. persimilis* females are present cannot be rejected. Because of small numbers of heterogamic matings, similar experiments with two strains of *D. prosaltans* furnish little evidence on this point. For *D. persimilis-D. pseudoobscura*, heterogamic matings are significantly more frequent in vials containing 10 females of each of the two species than in vials containing 10 alien females only, while for *D. prosaltans* they are significantly less frequent.

* Experimental data by Th. Dobzhansky, mathematical analysis by H. Levene.

[1] Mayr, E., and Dobzhansky, Th., these PROCEEDINGS, **31**, 75–82 (1945).

[2] For simplicity, we shall use the terms "conspecific" and "alien" regardless of whether the flies differ as to species (*D. persimilis* and *D. pseudoobscura*) or merely as to strain (geographic strains of *D. prosaltans*).

[3] Stalker, H. D., *Genetics*, **27**, 238–257 (1942).

[4] Throughout this paper Greek letters are used for population values (i.e., values characteristic of a given experimental set-up) and the corresponding Roman letters are used for estimates based on the sample of flies observed. Thus the best estimate of $\pi_{\sigma\delta}$ is $p_{\sigma\delta} = F_{\sigma\delta}/N_{\sigma\delta}$.

[5] Fisher, R. A., *Statistical Methods for Research Workers*, London, Sec. 53 ff., 2nd (1928) and later editions.

[6] Neyman, J., and Pearson, E. S., *Biometrika*, **20A**, 175–240 and 263–294 (1928).

[7] Wald, A., *Trans. Am. Math. Soc.*, **54**, 428–482 (1943).

[8] Dobzhansky, Th., and Streisinger, G., these PROCEEDINGS, **30**, 340–345 (1944).

THE THREE-DIMENSIONAL SHAPES OF BUBBLES IN FOAMS

BY EDWIN B. MATZKE

DEPARTMENT OF BOTANY, COLUMBIA UNIVERSITY

Communicated July 23, 1945

The three-dimensional shapes of bubbles in foams have been discussed by scientists in diverse fields for nearly three centuries, but apparently exact studies have not heretofore been made.

Even in his *Micrographia*, published in 1665, Robert Hooke,[1] the first

describer of cells, alluded to "frothy bodies" and "congeries of very small bubbles," while his contemporary, Nehemiah Grew,[2] in 1682 drew more direct comparisons between plant cells and the "Froth of Beer or Eggs." Ever since, the similarity between cells in tissues and bubbles in foam has been stressed. Thus, Holman, as reported by Child,[3] made a foam which he compared with cells in masses, while Robert[4] published a striking series of photographs of groups of bubbles simulating the early cleavage stages of the gasteropod *Trochus*. Berthold[5] and Errera[6] both showed how cells and soap films have certain features in common. However, until Lewis[7, 8, 9] began his painstaking researches on the three-dimensional shapes of cells in tissues, no precise information was available on the shapes of either cells or bubbles in foam, even though they had often been compared.

If biologists have been more interested than other scientists in this topic, they are by no means the only ones who have discussed it. Lord Kelvin,[10, 11] the distinguished mathematician and physicist, who was strongly influenced by the scholarly volumes on soap films written by the blind Belgian physicist Plateau,[12] published several essays on "the division of space with minimum partitional area," "homogeneously," "into equal and similar parts," "all sameways oriented" and under conditions of "stable equilibrium." Kelvin solved the problem of division of space into equal and similar units with stable angles and each with maximum volume and minimum surface by describing a 14-faced figure, the minimal tetrakaidecahedron, with 8 undulating hexagonal faces and 6 quadrilateral faces with the sides curved. This figure has stable angles. A similar plane-faced tetrakaidecahedron, with 8 faces which are regular hexagons and 6 which are squares, does not have completely stable angles. Kelvin's figures have been widely accepted, and under ideal conditions with *perfect spacing*, they may well be realized.

Sir James Dewar[13, 14, 15] made extensive studies of soap films and bubbles. Separate volumes have been written on the subject by Lawrence[16] and by Boys,[17] while Courant and Robbins[18] stress the importance of soap-film studies in the solution of certain mathematical problems.

But the only one who has seriously tried to examine the three-dimensional shape of bubbles in an actual foam is Desch,[19] who became interested in the problem through its bearing on the form of crystal grains in solidifying metals, as in marine propellers and in crucible steel. Desch's data are not extensive, and are somewhat at variance with the results presented below.

This same topic has been of interest also to students of the alveolar structure of protoplasm. Bütschli[20] concluded that protoplasmic alveoli were dodecahedral, while Seifriz[21] thinks they may be 12-faceted or 14-faceted.

It is obvious, therefore, that the three-dimensional shape of bubbles in foam has been of interest and of philosophical import to physicists, mathe-

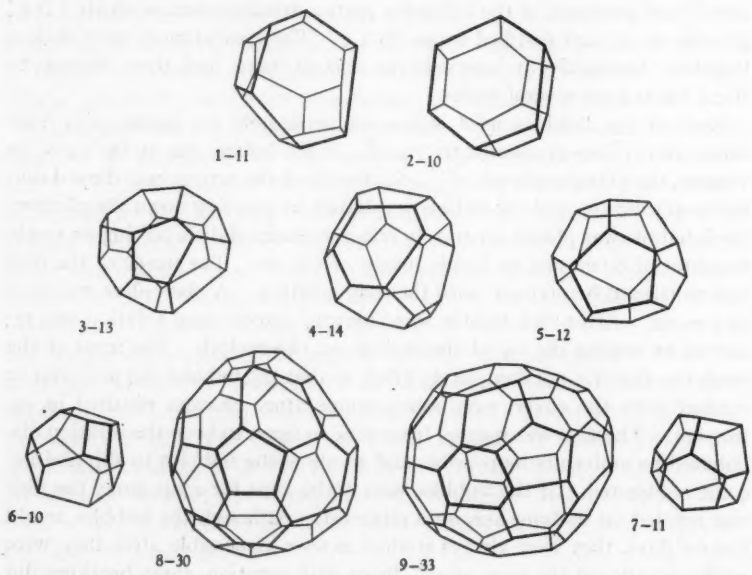
maticians, metallurgists, cellular biologists and students of protoplasmic structure. It is also evident that no one heretofore has succeeded in looking into a large mass of soap bubbles and tabulating the number and kinds of faces on each.

In order to accomplish this, at least under one set of conditions, a series of studies on the three-dimensional shapes of bubbles in foams have been carried out, using a technique different from that of previous investigators. The soap solution found most satisfactory was essentially that of Johnston,[22] and consisted of the following parts: triethanolamine oleate 7.5 g., glycerin 34 g., and distilled water 58.5 g. These substances were shaken together thoroughly at intervals for several days, and then allowed to stand for at least several weeks.

Each of the bubbles used was made separately by means of a Yale tuberculin syringe graduated to $^1/_{100}$ cc. If the bubble was to be $^1/_{10}$ cc. in volume, the plunger was set at $^1/_{10}$ cc., the tip of the syringe was dipped into the soap solution, and the bubble was blown by pressing down the plunger. Each bubble was placed separately into a cylindrical dish, having an inside diameter of 6 cm. and an inside height of 6.5 cm. The inside of the dish was moistened beforehand with the soap solution. A glass plate was used as a cover. After each bubble was "blown," excess soap solution was removed by wiping the tip of the syringe on cheesecloth. For most of the work the dish was not completely filled, so that the bubbles did not come in contact with the cover, even when temperature changes resulted in expansion. The dish was rotated from time to time, to keep the solution distributed as uniformly as possible and to allow the bubbles to slip and become readjusted. If the bubbles were to be kept for some time, the dish was rotated on its long axis on a clinostat. Although the bubbles would last for days, they were always studied as soon as possible after they were made, usually on the same day. Even with rotation, some breaking did occur, so that the mass of foam cells was not completely static, though a state of relatively stable equilibrium was achieved. When bubbles of varying sizes were used, the large ones tended to increase in volume, and the small ones to decrease, over a period of days. If breaking occurred, all the bubbles in the dish were discarded. Although only 1300 bubbles were used in the tabulations below, a total of approximately 57,200 bubbles were made, each individually, and placed separately into the dish in the course of the following experiments.

The bubbles, in the dish, were examined under a Spencer binocular dissecting microscope with 1X paired objectives and 6X paired oculars. If the films were thin, it was possible to focus through a dish of these bubbles and with the dissecting microscope to single out any individual one in a mass of approximately 1800, and to record exactly the number and kinds of faces.

The illustrations shown in figures 1 to 9 were made with a Leitz camera lucida attached to the dissecting microscope. The drawings were done by plotting the corners of each face and connecting them with straight lines. In this way the beautiful, delicate curvatures of the faces and edges, by which stability of angles is achieved, are not shown; this method was adopted for expediency, since to show these curvatures would be exceedingly difficult; the numbers and kinds of faces are truly represented.



FIGURES 1–9

Camera Lucida drawings of bubbles. Xc 3. The second number under each figure represents the total number of faces. Curvatures of faces and edges are not shown in the drawings. Quadrilateral faces are designated as Q, pentagonal as P, hexagonal as Hx, heptagonal as Hp. The combinations of faces in each are as follows: Fig. 1.—11 faces, 3 Q, 6 P, 2 Hx. Fig. 2.—10 faces, 4 Q, 4 P, 2 Hx. Fig. 3.—13 faces, 1 Q, 10 P, 2 Hx. Fig. 4.—14 faces, 2 Q, 8 P, 4 Hx. Fig. 5.—12 faces, 12 P. Fig. 6.—10 faces, 3 Q, 6 P, 1 Hx. Fig. 7.—11 faces, 2 Q, 8 P, 1 Hx. Fig. 8.—30 faces, 13 P, 16 Hx, 1 Hp. Fig. 9.—33 faces, 14 P, 17 Hx, 2 Hp.

In the first set of experiments, bubbles of uniform volume were used. The dish was filled, nearly to the top, with bubbles of $^1/_{10}$ cc. or $^1/_{20}$ cc. With each filling, bubbles of one volume only were utilized—they were not mixed. Since almost identical results were obtained with volumes of $^1/_{10}$ cc. and $^1/_{20}$ cc., all the data are combined in table 1.

It is immediately evident from table 1 that 400 peripheral bubbles—including 200 in contact with the walls of the cylinder and 200 on the free upper surface—had an average of 10.99 faces, and that the range was from 7 to 14. Bubbles with 11 facets occurred more commonly than any other kind. In table 1 it can be seen also that there were more pentagonal faces (2338) in these peripheral bubbles than all other kinds combined (2057).

TABLE 1

NUMBER AND KINDS OF FACES OF PERIPHERAL AND CENTRAL BUBBLES OF UNIFORM VOLUME

NO. OF FACES	NO. OF BUBBLES		KINDS OF FACES	PERIPHERAL	CENTRAL
	PERIPHERAL	CENTRAL			
7	2	...	Triangular	11	..
8	13	...	Quadrilateral	1252	866
9	20	...	Pentagonal	2338	5503
10	88	...	Hexagonal	720	1817
11	154	2	Heptagonal	73	35
12	83	73	Octagonal	1	..
13	33	179	Totals	4395	8221
14	7	218			
15	...	106			
16	...	20			
17	...	2			
	400	600			
Average	10.99	13.70			

The 400 peripheral bubbles here recorded showed 43 different combinations of faces. The combination occurring with greatest frequency had 3 quadrilateral, 6 pentagonal and 2 hexagonal faces (3 Q, 6 P, 2 Hx). It was found 67 times among the 400 peripheral bubbles tabulated, and is shown in figure 1. Normally the outer face was hexagonal, and there were 6 lateral and 4 basal contacts. The second most common combination in the peripheral bubbles (4 Q, 4 P, 2 Hx) is shown in figure 2; it occurred 47 times. Although certain of these combinations occurred much more frequently than others—9 of the 43 were found only once—still no one of these may be considered as "the type."

In addition to those on the surface, 600 "central" bubbles were examined, and the data are recorded in table 1. By "central" are meant those that were separated from the top, bottom or sides of the cylinder by at least three others. An inspection of the table reveals that the number of faces of the central bubbles was higher than that of the peripheral, that those with 14 contacts were most numerous, that the range was from 11 to 17,

and that the average for 600 was 13.70. It is further obvious that a great majority of the faces (5503 of 8221) were pentagonal.

The 600 central bubbles here tabulated were found in 36 different combinations of faces. The combination found most frequently was a 13-faceted foam cell with 1 quadrilateral, 10 pentagonal and 2 hexagonal contacts. It occurred 118 times among the 600, and is illustrated in figure 3. The second and third most common combinations—1 Q, 10 P, 3 Hx, and 2 Q, 8 P, 4 Hx (Fig. 4)—were both 14-faceted foam cells, and they were found 73 and 64 times, respectively. The fourth most common combination of the 36 was the pentagonal dodecahedron (Fig. 5) with 12 faces all pentagons; there were 50 of these. But, as in the peripheral bubbles, no one of these combinations may be singled out as "the type," and although the average number of contacts is close to 14, Kelvin's tetrakaidecahedra, with 8 hexagonal and 6 quadrilateral faces, are not realized under the conditions present here.

In a second set of experiments, bubbles of two different volumes were mixed together in known proportions. Small bubbles were made with a volume of 0.05 cc., and large bubbles with a volume of 0.4 cc. The results are summarized in table 2. In experiment 1, table 2, 2 small bubbles were made for each large one, so that when the dish was nearly filled and ready for examination, it contained twice as many small bubbles as large ones. In experiment 2, table 2, 8 small bubbles were put into the dish for each large one; in experiment 3, table 2, the ratio was 32 small to 1 large. Fifty small bubbles and 50 large ones were then studied in each of these 3 experiments, and the data recorded. Only "central" small and large bubbles were used—that is, none was tabulated unless it was at least the fourth one from the top, bottom or sides of the dish. During the course of these three experiments, in which 150 small and 150 large bubbles were studied, the dish was filled 19 times.

An analysis of table 2 reveals that when small and large bubbles are associated together, the number of contacts of the small ones is less than that of bubbles of uniform volume (13.70), while the number of contacts of the large ones is increased. Thus in experiment 1, table 2, the average of the small bubbles was 9.68, while that of the large ones was 20.42. Furthermore, increasing the ratio of small to large bubbles brought the number of contacts on the small ones closer to 13.70, and as the ratio of small to large increased, the number of contacts on the large bubbles became higher, within the limits of these experiments. On the basis of the data in table 2 and of the ratios of small to large bubbles used in the three experiments (that is, 2:1, 8:1, and 32:1), the average number of contacts of all the central bubbles (large and small) in the dish at one time can be computed; it is found to be 13.26 for experiment 1, 13.62 for experiment 2 and 13.71 for experiment 3. Therefore, when bubbles of two different volumes are asso-

ciated in varying but known proportions, the number of contacts is altered in definite ways.

TABLE 2

NUMBER AND KINDS OF FACES OF CENTRAL BUBBLES OF TWO DIFFERENT VOLUMES ASSOCIATED IN VARYING PROPORTIONS

	NUMBER OF BUBBLES						
NO. OF FACES	SMALL BUBBLES EXPT. 1	SMALL BUBBLES EXPT. 2	SMALL BUBBLES EXPT. 3	LARGE BUBBLES EXPT. 1	LARGE BUBBLES EXPT. 2	LARGE BUBBLES EXPT. 3	TOTALS
7	1	...	...	...	...	...	1
8	4	...	...	...	...	...	4
9	17	1	...	...	...	...	18
10	18	3	...	...	...	...	21
11	8	14	1	...	...	...	23
12	2	18	16	...	...	...	36
13	..	10	16	...	...	...	26
14	..	4	9	...	...	...	13
15	..	...	6	...	...	...	6
16	..	...	1	...	...	...	1
17	..	...	1	1	...	...	2
18	..	...	...	5	...	...	5
19	..	...	...	6	...	...	6
20	..	...	...	14	...	...	14
21	..	...	...	12	...	...	12
22	..	...	...	10	...	...	10
23	..	...	...	0	...	...	0
24	..	...	...	2	4	...	6
25	..	...	...	...	7	...	7
26	..	...	...	...	6	...	6
27	..	...	...	...	10	2	12
28	..	...	...	...	8	5	13
29	..	...	...	...	5	10	15
30	..	...	...	...	9	13	22
31	..	...	...	...	1	12	13
32	..	...	...	...	...	7	7
33	..	...	...	...	...	1	1
	50	50	50	50	50	50	300
Average	9.68	11.90	13.20	20.42	27.34	30.06	

	SMALL BUBBLES			LARGE BUBBLES			
KINDS OF FACES	EXPT. 1	EXPT. 2	EXPT. 3	EXPT. 1	EXPT. 2	EXPT. 3	TOTALS
Quadrilateral	159	118	96	115	57	36	581
Pentagonal	282	370	412	498	625	624	2811
Hexagonal	43	101	148	288	557	749	1886
Heptagonal	...	6	4	113	118	92	333
Octagonal	...	...	...	6	9	2	17
Nonagonal	...	...	...	1	1	..	2
	484	595	660	1021	1367	1503	5630

From table 2 it can be seen that pentagonal faces occurred more commonly than any other kind, constituting very nearly half the total number. Hexagons were also fairly abundant.

Varying the volumes and ratios of the bubbles results in an increase in the number of combinations of faces. Thus the 150 small bubbles were found in 35 different combinations, while the 150 large ones occurred in 105 combinations. Among the small bubbles the combination found most often (16 times) had 10 faces, 3 Q, 6 P, 1 Hx; it is illustrated in figure 6. Of the two second most abundant, both of which occurred 12 times, one had 11 faces, 2 Q, 8 P, 1 Hx, figure 7, and the other had 13 faces, 1 Q, 10 P, 2 Hx. This latter is also the commonest combination in central bubbles of uniform volume (Fig. 3). The large foam cell occurring most frequently (7 times) is shown in figure 8; it had 30 faces, 13 P, 16 Hx, 1 Hp. One bubble with 33 faces was found—the largest number recorded—14 P, 17 Hx, 2 Hp (Fig. 9). Here again, no one "type" is present in either the large or the small bubbles. Pentagonal faces occur frequently, hexagonal somewhat less so, but in no uniform pattern.

A consideration of the data presented in tables 1 and 2 leads to certain definite conclusions:

1. The peripheral foam cells of a mass of bubbles of uniform volume have an average of 11 faces. Lewis[23] had previously recorded 11-faceted epidermal cells.

2. The central bubbles in a foam, in which all are of approximately equal volume, have an average of nearly 14 faces (13.70). But pentagonal faces predominate, so that Kelvin's tetrakaidecahedra with 8 hexagonal and 6 quadrilateral faces are not realized under these conditions. This does not mean that Kelvin was incorrect, for he was interested in the theoretical division of space homogeneously into equal and similar units, similarly oriented and having stable angles. When a large number of bubbles of equal volume are placed into a cylindrical dish, so that they are free to glide and adjust themselves, they become arranged in a mass of somewhat irregular foam cells, approximating, at least, the equilibrium conditions outlined by Plateau, but Kelvin's "ideal" figure is not achieved. This is at variance with the conclusion presented in Sir D'Arcy Thompson's[24] scholarly volume.

3. Comparisons may be made between the data presented above and similar data on lead shot of uniform volume (Marvin[25]) and of varying volumes (Matzke[26]) compressed to eliminate interstices. In such a system, starting with spherical shot, the effects of surface energy in subsequent changes in shape are obviously of minor importance, in contrast to the conditions prevailing in soap films. The range in number of contacts, in kinds of faces, and especially in number of combinations of faces is more restricted in the foam cells than in the compressed lead shot.

4. Detailed comparisons will be published elsewhere between the complete data on soap films and cells in plant and animal tissues, described by Lewis, Marvin,[27, 28] Higinbotham,[29] Dodd[30] and Hulbary.[31] It may be

noted here, however, that cells in essentially undifferentiated tissues are intermediate in many respects between bubbles in foam on the one hand and compressed lead shot on the other. And frequently the cells approach the bubbles in foam distinctly more closely than they do the compressed lead shot. This is interpreted as indicating that surface forces are operative when the delicate cell walls are laid down in or near the plant meristems, that surface forces are of relatively less moment in cellular organization than in soap films, and that they constitute one of the factors, but only one, in three-dimensional cell shape determination.

[1] Hooke, R., *Micrographia*, London, 1665.
[2] Grew, N., *The Anatomy of Plants*, 2nd ed., London, 1682.
[3] Child, J. M., *Jour. Franklin Inst.*, **80**, 252–253 (1880).
[4] Robert, A., *Arch. de Zoologie Exp. et Gén.*, 3 ser., **10**, 269–538 (1902).
[5] Berthold, G., *Studien über Protoplasmamechanik*, Leipzig, 1886.
[6] Errera, L., *Bot. Centralbl.*, **34**, 395–398 (1888).
[7] Lewis, F. T., *Proc. Amer. Acad. Arts and Sci.*, **58**, 537–552 (1923).
[8] Lewis, F. T., *Ibid.*, **61**, 1–34 (1925).
[9] Lewis, F. T., *Ibid.*, **68**, 251–284 (1933).
[10] Kelvin, Lord, *Phil. Mag.*, 5s, **24**, 503–514 (1887).
[11] Kelvin, Lord, *Proc. Roy. Soc. London*, **55**, 1–16 (1894).
[12] Plateau, J., *Statique expérimentale et théorique des liquides soumis aux seules forces moléculaires*, Paris, 1873.
[13] Dewar, J., *Collected Papers of Sir James Dewar*, Vol. 2, 1154–1176 (1916).
[14] Dewar, J., *Collected Papers of Sir James Dewar*, Vol. 2, 1176–1206 (1917).
[15] Dewar, J., *Collected Papers of Sir James Dewar*, Vol. 2, 1206–1241 (1918).
[16] Lawrence, A. S. C., *Soap Films*, London, 1929.
[17] Boys, C. V., *Soap-bubbles and the Forces Which Mould Them*, London, 1890.
[18] Courant, R., and Robbins, H., *What Is Mathematics?* New York, 1941.
[19] Desch, C. H., *Jour. Inst. Metals*, **22**, 241–263 (1919).
[20] Bütschli, O., *Investigations on Microscopic Foams and on Protoplasm.* Trans. by E. A. Minchin, London, 1894.
[21] Seifriz, W., *Protoplasm*, New York, 1936.
[22] Johnston, B., *Civil Eng.*, **5**, 698–701 (1935).
[23] Lewis, F. T., *Science*, **86**, 609–611 (1937).
[24] Thompson, D'Arcy, *On Growth and Form*, New ed., 1942, p. 552.
[25] Marvin, J. W., *Amer. Jour. Bot.*, **26**, 280–288 (1939).
[26] Matzke, E. B., *Ibid.*, **26**, 288–295 (1939).
[27] Marvin, J. W., *Ibid.*, **26**, 487–504 (1939).
[28] Marvin, J. W., *Ibid.*, **31**, 208–218 (1944).
[29] Higinbotham, N., *Ibid.*, **29**, 851–858 (1942).
[30] Dodd, J. D., *Ibid.*, **31**, 120–127 (1944).
[31] Hulbary, R. L., *Ibid.*, **31**, 561–580 (1944).

GENETIC ANALYSIS OF THE INDUCTION OF TUMORS BY METHYLCHOLANTHRENE. XI. GERMINAL MUTATIONS AND OTHER SUDDEN BIOLOGICAL CHANGES FOLLOWING THE SUBCUTANEOUS INJECTION OF METHYLCHOLANTHRENE

BY LEONELL C. STRONG*

FROM THE DEPARTMENT OF ANATOMY, YALE UNIVERSITY SCHOOL OF MEDICINE

Communicated July 19, 1945

The injection into mice, or the application onto the skin surface, of methylcholanthrene or any one of a large number of carcinogens is followed by a great variety of neoplasms depending upon variations of experimental technique. The variation in the types of cancers obtained with the injection of methylcholanthrene has been considerably increased by the use of hybrid mice (from the F_2–F_{18} generations following an outcross) when the normally expected local tumors at the site of injection have been partially suppressed or completely inhibited through genetic selection toward resistance to such locally appearing tumors.[1]

The actual manner by which a carcinogen initiates normal tissues or cells to become cancerous is not clear. It is generally accepted, however, that the carcinogen or one of its metabolites acts directly upon cells in bringing about the neoplastic change. The observation of Mottram that carcinogens cause chromosomes to lag in the equatorial plate during mitosis is highly suggestive that nuclear changes may be involved in the origin of cancer.[2] The recent observation of the author that the untreated descendants of mice developing cancer of the stomach following the subcutaneous injection of methylcholanthrene continue to develop the same type of gastric cancer spontaneously is evidence that this particular neoplastic condition may have been hereditarily established through at least four generations by the effect of methylcholanthrene upon the germ plasm.[3] Linkage tests have demonstrated that susceptibility to gastric carconima is determined by the effect of a gene on the "brown tagged" chromosome.[3]

If it could be demonstrated that germinal changes (mutations) can be induced by methylcholanthrene, it would be strong presumptive evidence that perhaps the action of methylcholanthrene (and the other carcinogens) in bringing about cancer is the induction of a somatic mutation. This explanation is more consistent than the assumption that methylcholanthrene has one effect on the germ plasm (induction of mutations) and still another totally unrelated effect upon somatic tissue in converting them into cancer. Some evidence that mutations affecting susceptibility to tumors induced by methylcholanthrene do occur in mice following the injection of the carcinogen has already been published.[4]

It is the purpose of this paper to present evidence that germinal mutations

and other suddenly appearing biological characteristics arise in the descendants of mice that have received a subcutaneous injection of methylcholanthrene and that these sudden changes arise more frequently than are to be expected by chance alone.

Results.—One hundred fifty-five F_1 mice have been produced by a cross between C_{57} (homozygous for dark eye, black, non-agouti and self) and one subline of the NHO descent (proved to be homozygous for dark eye, brown, non-agouti and self). The above outcross was made in both directions. In each case only the NHO parent was injected with methylcholanthrene. Of these, 154 had the expected dark eye, black, non-agouti and self characters, while one was dark eye, black, non-agouti and spotted (proved to be dominant in inheritance). The new variant of spotting involved only the ventral surface of the body. Thus a dominant mutation of spotting occurred once in 155 mice. The author has not had dominant spotting (black eye white) in his laboratory for 25 years. It is also possible, although not definitely proved, that this new dominant spotting is not the same as the old black-eyed white character.

In all, thirteen color mutations have been obtained in the direct descendants of mice receiving methylcholanthrene. These are: (1) brown to black (dominant change), (2, 3, 4, 5) four pink eye mutants from dark eye (recessive inheritance), (6) color to albinism (recessive), (7) spotting involving only the ventral surface of the mouse discussed above (dominant), (8) a second type of spotting involving the nose and ventral surface only (dominant), (9) a third type of spotting involving only the belt region of the mouse (inheritance uncertain), (10) silvering (precocious graying evident upon the appearance of the first hair at 5–7 days of age) (inheritance uncertain), (11) permanent wave (inheritance uncertain), (12) piebald to self (dominant) and (13) gray belly to non-agouti in a form resembling black and tan. The thirteen new variants listed here are all proved to be *bona fide* mutations, based upon the observation of the subsequent appearance of the new forms in the lineal descendants of the mouse which first showed the character. The type of inheritance involved, determined by an analysis of the data obtained by an outcross to mice of an unrelated strain, has not been made in all cases—hence the use of the expression "type of inheritance uncertain."

Seven of these color mutations are repetitions of characters present in the author's laboratory, while six are new ones never having been present previous to the injection of methylcholanthrene.

The mutation rate is difficult to determine. For example, approximately 12,000 mice have been injected with 1 mg. of methylcholanthrene. Both parents of the NHO descent have been injected with the carcinogen for fourteen generations. The restrictions of space and time do not permit the keeping of all descendants of the experimental animals. In dealing with

hybrid mice as was done in this case, it is necessary to exclude all recessive variants that may have been due to segregation. A careful analysis of the data indicates that in these experimental animals in one restricted group the mutation rate was approximately 1 in 400[3]. In the controls over a period of 27 years, the mutation rate involving coat color changes has been approximately 1 in 26,000 mice.[3]

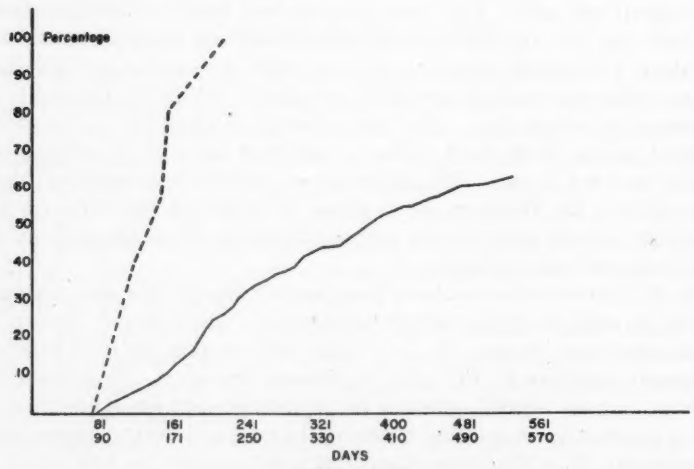


CHART ONE

Data on the altered susceptibility to local tumors induced by methylcholanthrene are given in this chart. Data on the original brown non-agouti subline are given on the solid line; data for the mutated black non-agouti subline are given on the dash line. Time in days is expressed along the base line; percentage incidence of induced local tumors is given on the vertical; that is, the data, with advancing time, are cumulative.

Two sudden changes, other than color changes and alterations affecting cancer susceptibility[4] have also arisen in the descendants of methylcholanthrene-treated animals. These are (1) the establishment of precocious sexual activity—the production of first litters by 42 days of age in one subline as compared to 78–85 days, the average in the original ancestral stocks which gave rise to them and (2) the establishment of a subline which has mice that give rise to extremely large first litters (average 10.2 young) as compared to 5.6 young in first litters of mice of the ancestral stocks.

Two other biological changes have also occurred. Whether they arose suddenly is not evident. These are (1) a more rapid growth rate in early life and (2) the attainment of adult body weight of 55–60 g. as compared to 26–28 g. for mice of the ancestral stocks.

The derived color mutant sublines either (1) retain the degree of cancer susceptibility (latent period measured from the time of injection to the origin of the local tumors) possessed by the ancestral stock mice that gave rise to them or (2) suddenly acquire a new susceptibility to induced tumors far in excess of the original suscpetibility (Chart 1). The impression is gained that many new biological characteristics are being acquired at approximately the same time. Thus the conclusion is reached that the effect of methylcholanthrene upon the germ plasm at any one time is either (1) the induction of a single gene mutation or (2) changes far in excess of what are to be expected by a single point mutation.

Summary.—Thirteen mutations involving coat color have appeared in the untreated descendants of mice receiving methylcholanthrene. The first mutant obtained occurred in a subline of the NHO strain which had been subjected to methylcholanthrene in both parents for twelve generations. An increased growth rate, the attainment of a larger body size and weight, precocious sexual activity, large first litters and increased susceptibility to induced local tumors by methylcholanthrene have also been obtained. These suddenly appearing color mutations have occurred either alone or in association with one or more of the changed biological characteristics enumerated above. Thus there is evidence that methylcholanthrene has affected the germ plasm by bringing about germinal or point mutations and perhaps other undetermined effects. It is highly probable, therefore, that methylcholanthrene may also bring about malignancies in tissues by causing somatic mutations to arise in them.

* This experiment has been made possible by grants from the Jane Coffin Childs Memorial Fund for Medical Research and the Anna Fuller Fund.

[1] Strong, L. C., "Genetic Nature of the Constitutional States of Cancer Susceptibility and Resistance in Mice and Men," *Yale Jour. Biol. and Med.*, **17** 289–299 (1944).

[2] Mottram, J. C., "Some Effects of Cancer-producing Agents on Chromosomes," *Brit. Jour. Exper. Path.*, **15** 71–73 (1934).

[3] Strong, L. C., "Genetic Analysis of the Induction of Tumors by Methylcholanthrene. IX. Induced and Spontaneous Adenocarcinomas of the Stomach in Mice," *Jour. Nat. Cancer Inst.*, **5** 339–362 (1945).

[4] Strong, L. C., "Genetic Analysis of the Induction of Tumors by Methylcholanthrene VIII. Two Mutations Arising in Mice Following Injection of Methylcholanthrene," *Arch. Path.*, **39** 232–236 (1945).

DAMPING OF EPIDEMIC WAVES

BY EDWIN B. WILSON AND JANE WORCESTER

HARVARD SCHOOL OF PUBLIC HEALTH

Communicated July 12, 1945

Soper[1] stated: "In the foregoing theory . . . infecting power was supposed to be instant, at the termination of a certain incubation period, and it appears that with such a law in operation, an initial upset of steady conditions of prevalence will be followed by epidemic waves that propagate themselves indefinitely, without diminution of amplitude." And he went on to develop another theory in which infecting power begins at the instant a person becomes infected and is spread over time but with a diminishing incidence; on this assumption he finds that there is damping of the epidemic waves. His arguments, be it noted, assume that the disturbance from the steady state may be treated as infinitesimal.

The first conclusion, namely, that, under the assumption that there is a fixed incubation period, there will be no damping, is incorrect, and its incorrectness derives from a particular one of the approximations Soper makes. He starts out with the equations

$$C(t) = (S/m)\, C(t - \tau), \qquad C = A - dS/dt \tag{1}$$

where C is the case rate, m a constant, S the number of susceptibles and A the rate of accession of susceptibles supposed steady, and he remarks that since the change in S is usually small in the unit interval (the incubation period τ) we may write these equations as

$$C(t + \tau/2) = (S/m)\, C(t - \tau/2), \qquad C = A - dS/dt. \tag{2}$$

This change in effect removes the lag due to the incubation period and therefrom is derived the absence of damping in the wave.

If we use D as a symbol of differentiation and $e^{hD}f$ as the symbol for Taylor's expansion and if we set $S = m + y$, with y small so that higher powers are to be neglected (1) and (2) become, respectively,

$$[\tau D(1 - e^{-\tau D}) + A\tau/m]y = 0, \tag{1'}$$

$$\left[\frac{\tau D}{2} \sinh \frac{\tau D}{2} + \frac{A\tau}{4m}\right] y = 0. \tag{2'}$$

The second equation has a negative root for D^2 and hence an undamped period in the neighborhood of $P = 2\pi\sqrt{m\tau/A}$. The first equation does not have a pure imaginary solution for D, but a solution with a negative real part leading to damped harmonic waves of the approximate form

$$y = e^{-At/4m} \cos \sqrt{A/m\tau}\, t$$

with period $P = 2\pi\sqrt{m\tau/A}$ (approximately[2]) and damping per period of $e^{-\pi/2\cdot\sqrt{A\tau/m}}$.

We may give the solution for a more general equation in which the generalized form of the law of mass-action is used and not only an incubation period τ but a duration σ of infectiveness (both constant) are assumed. The equation in S is then[3]

$$A - \left(\frac{dS}{dt}\right) = \left(\frac{S}{m}\right)^p [A + S(t - \tau - \sigma) - S(t - \tau)]$$

and if $S = m + y$ with y infinitesimal we have

$$\left[D + \frac{1}{\sigma} e^{-(\tau+\sigma)D} - \frac{1}{\sigma} e^{-\tau D} + \frac{Ap}{m}\right] y = 0.$$

If we let $a = \tau + \sigma/2$ we have

$$\left[aD^2 - \left(\frac{a^2}{2} + \frac{\sigma^2}{24}\right) D^3 + \left(\frac{a^3}{6} + \frac{a\sigma^2}{24}\right) D^4 - \ldots + \frac{Ap}{m}\right] y = 0.$$

When Ap/am is small the solution is approximately[4]

$$D = \frac{1}{4}\frac{pA}{m}\left(1 + \frac{\sigma^2}{12a^2}\right) + i\sqrt{\frac{pA}{ma}} \qquad (3)$$

and from this the period and damping can be read off.

For the case of Soper's distributed infectiousness we should write in our notation (with $p = 1$ in the law of mass-action)

$$A - \frac{dS}{dt} = \frac{S}{m}\int_{-\infty}^{t} \left(A - \frac{dS}{dt}\right)_\xi e^{(\xi - t)/\tau}\, d\xi$$

so that new cases at time $\xi < t$ are effective at time t only in proportion[5] $e^{(\xi-t)/\tau}$. The mean value of the duration of infectiveness is then

$$\int_{-\infty}^{0} \xi e^{\xi/\tau} \div \int_{-\infty}^{0} e^{\xi/\tau}\, d\xi = \tau.$$

Now it so happens that[6]

$$\int_{-\infty}^{t} e^{(\xi-t)/\tau} f(\xi) d\xi = \frac{1}{1 + \tau D} f.$$

Hence the equation with the generalized mass law would be

$$A - \frac{dS}{dt} = \left(\frac{S}{m}\right)^p \left[A - \frac{D}{1 + D} S\right]$$

and for $S = m + y$, with y infinitesimal,

$$\left[D - \frac{D}{1 + \tau D} + \frac{Ap}{m}\right] y = 0 \quad \text{or} \quad \left[D^2 + \frac{Ap}{m} D + \frac{Ap}{m\tau}\right] y = 0.$$

In this case the value of D may be obtained in finite form as

$$D = -\frac{Ap}{2m} + i\sqrt{\frac{Ap}{m\tau}}\sqrt{1 - \frac{Ap\tau}{4m}}.$$

If compared with (3) taking $\sigma = 0$ and $a = \tau$ we see that the waves have approximately the same period whether we use a fixed incubation period τ and an infinitesimal duration of infectiousness or whether we use an infinitesimal incubation period with a period of infectiousness which averages to be τ, it being understood that the infectiveness dies out exponentially. The former case has, however, half the damping of the latter instead of none at all as Soper states.

It is of further interest in connection with stepwise calculations to note that one may not replace derivatives in differential equations with differences without becoming liable to modify the indicated damping. For example, in illustrating his theory numerically Soper replaces

$$\frac{dx}{dt} = a - a\frac{x}{as}\frac{y}{a\tau}, \qquad \frac{dy}{dt} = a\frac{x}{as}\frac{y}{a\tau} - a\frac{y}{a\tau}$$

for purposes of calculation by

$$\Delta x = a - a\frac{x}{as}\frac{y}{a\tau}, \qquad \Delta y = a\frac{x}{as}\frac{y}{a\tau} - a\frac{y}{a\tau}.$$

If we set $x = as(1 + u)$, $y = a\tau(1 + v)$, with u and v infinitesimal the differential equations, of the infinitesimal epidemic (as found by Soper) are

$$s\frac{du}{dt} = -u - v, \quad \tau\frac{dv}{dt} = u,$$

the solution of which leads to the equation for D

$$\begin{vmatrix} sD + 1 & 1 \\ -1 & \tau D \end{vmatrix} = 0 = s\tau D^2 + \tau D + 1,$$

$$D = -\frac{1}{2s} + i\sqrt{\frac{1}{s\tau} - \frac{1}{4s^2}}.$$

If, however, we replace D by $\Delta = e^D - 1$ we have

$$\begin{vmatrix} s(e^D - 1) + 1 & 1 \\ -1 & \tau(e^D - 1) \end{vmatrix} = 0 = s\tau e^{2D} - (2s\tau - \tau)e^D + s\tau - \tau + 1,$$

of which the solution is

$$D = \frac{1}{2}\log_e\left(1 - \frac{1}{s} + \frac{1}{s\tau}\right) + i\tan^{-1}\frac{\sqrt{s/\tau - 1/4}}{s - 1/2}.$$

If $s = 68.2$ and $\tau = 2$ the values of the pure imaginary parts of the two expressions for D will be nearly the same and thus the periods will also be; but the real part in the second case will be about $-1/4s$ whereas in the first it is $-1/2s$, i.e., twice as much.[7] This explains why Soper found numerically about 0.80 instead of about 0.58.

We have therefore seen that the damping on the "instant hypothesis" is determined (approximately) by the logarithmic decrement $Ap/4m$ instead of zero as stated by Soper and by $Ap/2m$ on the "hypothesis of distributed infectiousness." We have further seen that the damping is modified by replacing derivatives by increments. The rough equations given in an earlier paper[3] for the stepwise calculation of an epidemic with recruits, when applied to a variety of hypothetical cases, appear to give no damping and, as the theory indicates damping, those equations presumably have actually eliminated the lag to which damping would be due. However, it must be admitted that the phenomenon of recurrent measles epidemics gives no clear evidence of any damping. This creates something of a difficulty with the theory in respect to the prediction of damping and throws some doubt on the reality of periods; it is possible that measles simply dies out and then returns and under such a hypothesis there would seem to be no reason to expect either definite periods or damping to be observable by comparing successive epidemics.

[1] Soper, H. E., "The Interpretation of Periodicity in Disease Prevalence," *Jour. Roy. Statist. Soc., London*, **92**, 34–61 (1929).

[2] If $\alpha = A\tau/m$ and the root for τD be $X + iY$ we have to solve $X - Xe^{-X}\cos Y - Ye^{-X}\sin Y + \alpha = 0$, $Y - Ye^{-X}\cos Y + Xe^{-X}\sin Y = 0$ and the solutions are of the form $X = -1/4\,\alpha\,[1 + c_1\alpha + c_2\alpha^2 + \ldots]$, $Y = \alpha^{1/2}\,[1 + d_1\alpha + d_2\alpha^2 + \ldots]$. The value of α for measles is a small number; if m be $5\frac{1}{2}$ years of recruits to the susceptible population and τ be half a month, $\alpha = 1/132$.

[3] See these PROCEEDINGS, **28**, 361–367 (1942) and **31**, 24–34, 109–116 (1945).

[4] The value of aD may be written as a power series in pAa/m. If we set

$$R^2 = \frac{pAa}{m},\ L = \frac{1}{2} + \frac{\sigma^2}{24a^2},\ M = \frac{1}{6} + \frac{\sigma^2}{24a^2},\ N = \frac{1}{24} + \frac{\sigma^2}{48a^2} + \frac{\sigma^4}{1920a^4},$$

$$AD = iR\left[1 + \left(\frac{1}{2}M - \frac{5}{8}L^2\right)R^2\right] - \frac{1}{2}LR^2 + \left(\frac{1}{2}N - \frac{3}{2}LM + L^3\right)R^4$$

this second approximation will give an estimate of the accuracy of the first approximation when definite values of σ, τ, p, A, m are assumed. As $a = \tau + \sigma/2$ we must have σ/a varying from 0 to 2 as σ/τ varies from 0 to ∞. Soper really assumed $\sigma = 0$, $\tau = a$ fortnights for measles.

[5] The new-case-rate (of becoming infectious) at time t, viz., $A - dS/dt$, can be regarded as the sum of the products of S by the product

$$\frac{1}{m}\left(A - \frac{dS}{dt}\right)_\xi e^{(\xi - t)/\tau}\,d\xi = \frac{d\xi}{m}\left[\left(A - \frac{dS}{dt}\right)_\xi e^{(\xi - t)/\tau}\right] = \left(A - \frac{dS}{dt}\right)_\xi\left[\frac{1}{m}e^{(\xi - t)/\tau}\right]d\xi$$

contributed by the cases which become infectious at the time ξ previous to t, where the

contact rate is $r = 1/m$. The expression is capable of two interpretations, viz., (*a*) the contact rate remains constant but of the number of cases which become infectious at the time ξ only the fraction $e^{(\xi-t)/\tau}$ remain infectious, which seems to be Soper's hypothesis and which accords with the way in which the carrier condition in diphtheria falls off with the time, at least after the initial infectiveness (see Stallybrass, C. O., *Principles of Epidemiology*, pp. 302 and 617); (*b*) the cases which become infectious at time ξ all remain infectious but the contact rate r falls off as $e^{(\xi-t)/\tau}$. In the first interpretation τ is the average period of infectiousness, in the second it is the period required for the infectiousness to fall off to the fraction $1/e$ of its initial value; in the first we should speak of the diminishing incidence of the infectiousness, in the second of its diminishing intensity.

[6] See, for example, Wilson, E. B., *Advanced Calculus*, p. 451, Exs. 17, 18.

[7] If the root for D is $-X + iY$, the damping in time is e^{-Xt} and if X be divided by 2 the damping in the same time, being $e^{-Xt/2}$, is the square root of e^{-Xt}. There is, however, a further distinction which must be borne in mind. The differential equations are in terms of $u = \log(C/A)$. If C/A is near 1 so that the epidemic is infinitesimal, it is not the case-rate C which is directly subject to damping, but the part of it which exceeds A as $\log(C/A)$ is practically equal to $(C - A)/A$. If C is really near to A, the value of C is practically not damped appreciably. With $A = 2200$, $C_0 = 6600$, Soper's calculation gave him a second peak of $C_1 = 5168$ which is the fraction 0.783 of the first, and may be taken as a sort of damping factor—very likely the sort an epidemiologist would think most natural. However, to compare with the theory of the infinitesimal epidemic one would have to take $u_0 = \log(6600/2200) = 1.099$ and $u_1 = \log(5168/2200) = 0.8544$. The ratio of 0.8544 to 1.099 is 0.777 which is very near to 0.783. The theoretical value for the infinitesimal epidemic is 0.584 but when $\Delta = e^D - 1$ replaces D as an operator the theoretical solution gives, not 0.584, but 0.764 which certainly is as near as could be expected to the value 0.777 Soper found.

Of course, an epidemic with $C_0/A = 3$, $u_0 = 1.099$ can hardly be considered infinitesimal; the period will doubtless have increased somewhat (as Soper actually found in his calculation) and the theoretical damping will presumably have been somewhat modified. It should be noted that with $u_0 = 1.099$ we have, if we neglect effects of damping, from the equation $u_1 - e^{u_1} + e^{u_0} - u_0 = 0$ between the trough u_1 and crest u_0, the value $u_1 = -1.72$; the values $+1.10$ and -1.72 are so far from symmetrically situated with respect to $u = 0$ as to indicate a major departure from the simple harmonic deviation of the infinitesimal oscillation. Indeed it can no longer be expected that the damping as estimated by the fraction that one crest is of the preceding one and the fraction that one trough is of the preceding one should be the same. A satisfactory approximate solution of our equation (3) of a previous paper[8]

$$\left.\frac{du}{dt}\right|_{t+\tau/2} - \left.\frac{du}{dt}\right|_{t-\tau/2} = \frac{pA}{m} e^{u(t-\tau/2) - u(t+\tau/2)/p} \left(1 - e^{u(t+\tau/2)}\right)$$

would be required probably to discuss periodicity and damping in the case of finite epidemics.

TOPOLOGICAL METHODS IN THE THEORY OF FUNCTIONS OF A SINGLE COMPLEX VARIABLE. I. DEFORMATION TYPES OF LOCALLY SIMPLE PLANE CURVES*

BY MARSTON MORSE AND MAURICE HEINS

INSTITUTE FOR ADVANCED STUDY

Communicated July 25, 1945

1. Canonical Curves.—Title I refers to the first of a series of four papers by the authors under the general heading of "Topological Methods in the Theory of Functions of a Single Complex Variable." An abstract of Paper II will follow this abstract.

In the study of meromorphic functions one comes most naturally to *locally simple* sensed closed plane curves and deformations of these curves through families of sensed curves which are *uniformly* locally simple. A closed curve g is *locally simple* if for some sufficiently small positive constant e each point P of g is an interior point of a simple subarc of g whose endpoints P_1 and P_2 satisfy the condition

$$|P_1P| > e \qquad |P_2P| > e.$$

A set of closed curves whose points P satisfy the preceding conditions for one and the same e will be termed *uniformly* locally simple. A curve may have infinitely many multiple points and be locally simple, and there are continuous families of locally simple closed curves which are not uniformly locally simple.

We admit deformations through continuous families of uniformly locally simple closed curves.

A first theorem is that any locally simple sensed closed plane curve g can be admissibly deformed into a curve C^n obtained by tracing a sensed circle

$$x = \cos\theta \qquad y = \sin\theta \tag{1.1}$$

n times, $n = \pm 1, \pm 2, \ldots$ or into a figure eight, C°. Moreover g can be deformed into but one of these canonical curves.

In particular the figure eight, C°, traced m times can be admissibly deformed into C°.

If a locally simple curve is given as the locally 1–1 continuous image of the circle (1.1) in the form

$$w = f(\theta) \qquad (w = u + iv)$$

then as θ increases from 0 to 2π any continuous branch of the function

$$\text{arc}\,[f(\theta + e) - f(\theta)] \qquad (0 < e \leqq e_1)$$

will change by an integral multiple $2p\pi$ of 2π, independent of e provided e_1 is a sufficiently small positive constant. The integer p is called the *angular order* of g.

It is apparent that the angular order of an admissible curve g is invariant under any admissible deformation of g. It is shown that an admissible curve of angular order p can be admissibly deformed into the above canonical curve C^p, including the figure eight C° when $p = 0$.

A polygon π is locally simple if no edge has more than one end-point in common with the succeeding edge of the polygon. A continuous variation of a circular sequence of vertices representing a polygon will define an admissible deformation of π if successive vertices are bounded from each other and if successive edges have at most an end-point in common. We admit the addition or removal of a finite number of vertices on edges of a polygon at a finite number of moments t of a deformation.

The first approach to the preceding theorems is by way of lemmas by virtue of which any locally simple sensed curve can be admissibly deformed into an admissible polygon π. It is then shown that π can be admissibly polygonally deformed either into a triangle traced n times, $n = \pm 1, \pm 2, \ldots$, or into two triangles with a vertex in common forming a figure eight. The proof is by induction with respect to the number of multiple points.

2. Products of Deformation Types.—The class of admissible curves admissibly deformable into an admissible curve g will be denoted by (g) and termed a deformation type T. Each deformation type includes a sensed regular curve. With the aid of such regular curves the *product*

$$T_1T_2 = (g_1)(g_2)$$

will be defined. Let h_1 and h_2 be regular sensed curves in (g_1) and (g_2), respectively, which are positively tangent at a point P. Let h_1h_2 be a closed curve obtained by cutting h_1 and h_2 at P, to form arcs h' and h'', respectively, tracing h', then h'', and joining the final end-point of h'' to the initial end-point of h'. We define T_1T_2 as the deformation type (h_1h_2) and show that this type depends only on T_1 and T_2, and not on P or the choice of h_1 and h_2 in the types T_1 and T_2.

It is shown that the product T_1T_2 is commutative and associative.

For n positive or negative let g^n denote g traced n times. For $n = 0$ let g^n denote a curve of the type of $(g)(g^{-1})$. With these definitions it is shown that

$$(g^r)(g^m) = (g^{r+m})$$

where r and m are arbitrary integers. It must be clearly understood that the operation $(\ )^m$ is *a priori* utterly different from the operation of tracing g m times.

It follows that the deformation types form an Abelian group with respect to the product operation, with the type of the figure eight the idem-factor. This group is isomorphic with the additive group of integers.

3. The Order q.—The order $q(w)$ of a sensed closed curve g with respect to point w not on g is well defined.

If g is locally simple, it is shown that the order $q(w)$ possesses an absolute bound independent of points w not on g.

An admissible deformation none of whose curves intersects a fixed point O will be called an *O-deformation*. It will be convenient to take O as the origin in the w-plane.

We need canonical curves with respect to O-deformations, given p and q. To that end let C be the sensed unit circle previously defined. Let C_1 be a similarly sensed circle of radius 1/4, tangent to C at $(u, v) = (1, 0)$. Let

$$C^q C_1^r \qquad (q \neq 0, r \neq 0) \tag{1.2}$$

denote the curve obtained by first cutting C and C_1 at (1, 0), starting with the point (1, 0), then tracing C q times in the sense indicated by the sign of q, next tracing C_1 similarly r times, closing the curve at (1, 0). The curve C_1 shall be taken *interiorly* or *exteriorly* tangent to C at (1, 0) according as q and r have the same or different signs. The resulting curve will be locally simple and regular. It is observed that the curve (1.2) has the order q with respect to $w = 0$ and the angular order p.

There exists an admissible curve g with prescribed order q and angular order p. A principal theorem follows:

Let g be an admissible curve of order q relative to the origin and of angular order p. Let the following cases be distinguished

Case I $q \neq 0$ $\quad p - q \neq 0$
Case II $q \neq 0$ $\quad p - q = 0$
Case III $q = 0$.

Then g admits an O-deformation into a curve $C^q C_1^{p-q}$ in Case I, a curve of the form C^p in Case II, and a curve of the form C_1^p in Case III. No one of these canonical curves with the orders (p, q) admits an O-deformation into a canonical curve with a different pair of orders.

The relation of this deformation theory to the theory of meromorphic functions will become clear in Paper II.

* Morse, M., and Heins, M., "Topological Methods in the Theory of Functions of a Single Complex Variable: II. Boundary Values and Integral Characteristics of Interior Transformations and Pseudo-harmonic Functions. III. Causal Isomorphisms in the Theory of Pseudo-harmonic Functions." Papers I and II are to appear in the *Annals of Mathematics*.

Morse, M., "The Topology of Pseudo-harmonic Functions." Completed but not yet submitted for publication. This paper obtains the basic topological relations A upon which Paper II is based. Paper III requires the use of infinite groups, and obtains relations A as a special case. Relations A can be obtained of themselves more simply and are so obtained in the present paper.

TOPOLOGICAL METHODS IN THE THEORY OF FUNCTIONS OF A COMPLEX VARIABLE. II. BOUNDARY VALUES AND INTEGRAL CHARACTERISTICS OF INTERIOR TRANSFORMATIONS AND PSEUDO-HARMONIC FUNCTIONS*

BY MARSTON MORSE AND MAURICE HEINS

INSTITUTE FOR ADVANCED STUDY

Communicated July 25, 1945

1. Interior Transformations.—The objective of the authors is to find out what properties of meromorphic functions are essentially topological in character, and by the introduction of appropriate topological concepts and methods to extend and simplify the present theory. To that end it is natural to replace the class of meromorphic functions by their topological generalizations, the interior transformations, studied by Stoilow,[1] Whyburn[2] and others, and the class of harmonic functions by a similar topological generalization termed by us pseudo-harmonic functions. The point of the present series of papers is by no means the generality obtained in this way but rather the insight and control of the theory, and the new problems and results which are brought out.

Let $w = F(t) \not\equiv$ const. be a transformation of a neighborhood N in the t-plane of a point t_0 into a neighborhood N_0 of w_0 (possibly the point at infinity) where $F(t)$ is analytic at t_0 or has a pole at t_0. Let $t = \varphi(z)$ map z_0 onto t_0 and N homeomorphically onto a neighborhood N_1 of z_0. Then

$$w = F(\varphi(z)) = f(z)$$

will be termed an *interior transformation* of N_1 into N_0. The point z_0 will be termed a zero or pole of $f(z)$ if t_0 is a zero or pole, respectively, of $F(t)$, and the order of z_0 as a zero or pole of $f(z)$ will be taken as the order of t_0 as a zero or pole of $F(t)$. If the inverse of $F(t)$ has a branch point of the mth order belonging to the pair (t_0, w_0) the inverse of $f(z)$ will be said to have a branch point of the mth order belonging to the pair (z_0, w_0).

We shall be concerned with a limited connected region G in the z-plane bounded by ν Jordan curves

$$(\beta) = (\beta_1, \ldots, \beta_\nu).$$

Without loss of generality these curves can be taken as circles. We shall *admit* transformations $w = f(z)$ of $\overline{G}$ (the closure of G) into the extended w-plane such that $w = f(z)$ is an interior transformation of some neighborhood of each point of G and is continuous on $\overline{G}$ at each point of (β). Such a function $f(z)$ will be called an interior transformation of G.

The "integral characteristics" of $f(z)$ are as follows: $n(0)$ = the number of zeros of $f(z)$ on G; $n(\infty)$ = the number of poles of $f(z)$ on G; μ = the

number of ramification points with antecedents on G of the inverse of $f(z)$. These points are counted with their multiplicities.

The *order* q of a sensed closed curve g in the w-plane with respect to $w = 0$ is well defined provided g does not pass through $w = 0$. If g is locally simple in the sense of Morse and Heins, I (these PROCEEDINGS, preceding article), the *angular order* p of g is well defined. A basic theorem follows.

I. *If $w = f(z)$ is meromorphic on G and continuous at each point of (β), if the images g_t of the positively sensed boundary curves β_t are locally simple and do not intersect the origin, then*

$$n(0) + n(\infty) - \mu \geqq 2 - \nu + \Sigma q - \Sigma p \qquad (1.1)$$

summing over the boundary images g_t.

Examples exist in which the equality in (1.1) is excluded. Stoilow[1, b] has considered the case of one boundary curve whose image is a Jordan curve and has mistakenly affirmed that

$$n(0) + n(\infty) - \mu = 1.$$

For the Stoilow case $p = q = 1$, so that the right member of (1.1) is 1 but, as stated, the inequality need not hold.

II. *If, however, one adds the hypothesis that the transformation $w = f(z)$ is 1–1 in some neighborhood relative to G of each boundary point of G, then (1.1) becomes an equality.*

The theorem with the equality has been proved by us for the case of interior transformations, the relation (1.1) in I only for meromorphic function. The proof of I depends upon the lemma that the antecedents of points of ramification cluster at no boundary curve β_t whose image g_t is locally simple, and this has been proved only when $w = f(z)$ is meromorphic.

A generalization of a theorem of Radó[3] is given among various applications.

2. *Pseudo-harmonic Functions.*—If $f(z)$ is an interior transformation of a neighborhood N of z_0 without pole at z_0, then $Rf(z) = U(x, y)$ will be said to be *pseudo-harmonic* on N. In case $f(z)$ has a zero or pole at z_0 but is not identically zero, then

$$\log|f(z)| = U(x, y) \qquad (2.1)$$

will be said to be pseudo-harmonic on N except for a *logarithmic pole.* Functions $U(x, y)$ will be admitted on $\overline{G}$ which are pseudo-harmonic on some neighborhood of each point of G except for logarithmic poles and which are continuous on $\overline{G}$ at points of (β).

Saddle points of $U(x, y)$ on G are similar to saddle points of harmonic functions. To treat the boundary values we begin with *Boundary Condition A* by virtue of which the function $U^{(\beta)}$ defined by U on (β) has at most

a finite number of points of extremum. It is seen that the locus $U = c$ at a given level c consists of a finite number of closed arcs, or arcs which tend to definite points of (β). Neighboring an interior point or a boundary point which is not isolated at the level c, the regions "below c" alternate with the regions "above c" in a well-defined manner. The *multiplicity* μ of a saddle point P is one less than the number of these "sectors" below c neighboring P. The interior or saddle points of positive multiplicity together with the points of relative minimum of U are called the *critical points* of U.

Set M = the number of logarithmic poles of U, S = the number of saddle points of U on G, s = the number of boundary saddle points, m = the number of points of relative minimum of U, counting the saddle points with their multiplicities. Then

$$M - S = 2 - \nu + s - m. \tag{2.2}$$

It is noted that $U(x, y)$ may be without partial derivatives both on G and $\bar{G}$, and that if these partial derivatives exist they may be null on (β). Under such conditions the classical alternatives to our evaluations, such as the Kronecker integral index, are undefined and therefore inapplicable.

Boundary Conditions B.—There is a special evaluation of $s - m$ on the right of (2.2) when U is of class C' on a neighborhood of (β) and has a non-null gradient λ on (β). These conditions are termed *Boundary Conditions B*, and do not imply that U is differentiable in general on G.

To proceed one covers the points of (β) at which λ is normal to (β) and entrant into G, by a finite set Ω of non-intersecting open arcs which exclude from their closures each point of (β) at which λ is normal to (β) and exeunt from G.

Let $\lambda(P)$ be the projection of λ onto the tangent to (β) at the point P on (β).

The Vector Index J of (β).—Let h be any arc of Ω. If h is a curve β_i an index 0 is assigned to h. If h is an arc of Ω not a curve β_i, an index 1 is assigned to h. To each end-point P of h at which $\lambda(P)$ is entrant relative to h an index 1 will be assigned. To other end-points 0 will be assigned.

The *vector index* J_i of U on β_i is defined as $E_0 - E_1$, where E_0 is the sum of the indices of end-points of arcs of Ω and E_1 is the sum of indices of the arcs of Ω. To (β) will be assigned the vector index

$$J = \Sigma J_i.$$

A first theorem is that J is independent of the choice of the covering Ω *among admissible coverings. Then one readily proves that*

$$M - S = 2 - \nu + J$$

for the most general pseudo-harmonic function satisfying Boundary Conditions B. In the special case in which Boundary Conditions A and B both hold, J equals the number of points of maximum of $U^{(\beta)}$ at which λ is entrant minus the number of minimum of $U^{(\beta)}$ at which λ is entrant.

3. Comparison Theorems for Harmonic Functions.—We begin with the case where $U(x, y)$ is harmonic on the unit circle $x^2 + y^2 \leqq 1$ and continuous on the boundary. If $u(r, \theta)$ represents U in polar coordinates, it is found that when the boundary values $u(1, \theta) = p(\theta)$ are of class C' and $p'(\theta)$ satisfies a Hölder condition, then $u_r(1, \theta)$ exists, and

$$u_r(1, \theta_0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{p(\theta_0 + \alpha) - p(\alpha)}{\cos \alpha} \right] d\alpha \qquad [\text{if } p'(\theta_0) = 0].$$

A boundary point θ is called *entrant* if $u_r(1, \theta) < 0$ otherwise *exeunt*. The importance of these distinctions has already been indicated.

A function $p(\theta)$ of the above sort will be termed *R-admissible* if it has at most a finite number of relative maxima and minima occurring at isolated points between which $p'(\theta)$ is never 0.

If $p(\theta)$ is R-admissible, has N absolute minima and N absolute maxima and no other extrema, then the harmonic function with boundary values $p(\theta)$ has just $N - 1$ saddle points on the region $r < 1$.

In any case the number S of saddle points on $r < 1$ is at most $N - 1$. If $p(\theta)$ is R-admissible and has $2N$ points of extremum and if the corresponding function $u(r, \theta)$ has $S < N - 1$ saddle points on $r < 1$, and no differential critical points on $r = 1$, it is possible to successively modify $p(\theta)$ on arbitrarily small neighborhoods of the extremum points so that $p(\theta)$ remains R-admissible and none of its extremum values and points are changed, while the number of saddle points of $u(r, \theta)$ on the region $r < 1$ assumes each integer from S to $N - 1$, inclusive.

If $p(\theta)$ is R-admissible, it is possible to admissibly modify $p(\theta)$ on an arbitrarily small neighborhood of any one of its points of relative minimum (maximum) so that the resulting harmonic function has no critical points on the region $r < 1$.

Various other comparison theorems are given for harmonic functions with and without logarithmic poles.

* Morse, M., and Heins, M., "Topological Methods in the Theory of Functions of a Single Complex Variable: I. Deformation Types of Locally Simple Plane Curves. II. The present paper. III. Causal Isomorphisms in the Theory of Pseudo-harmonic Functions." Papers I and II are to appear in the *Annals of Mathematics*. For an abstract of Paper I see these PROCEEDINGS, **31**, 299–301 (1945).

Morse, M., "The Topology of Pseudo-harmonic Functions." Not yet submitted for publication.

[1] Stoilow, S., (*a*) "Du caractère topologique d'un théorème sur les fonctions méromorphes," *C. R. de l'Acad. des Sci. de Paris*, **190**, 251–253 (1930). (*b*) *Leçons sur les principes topologiques de la théorie des fonctions analytiques*, Paris, 1938.

[2] Whyburn, G. T., *Analytic Topology*, New York, 1942.

[3] Radó, T., "Zur Theorie der mehrdeutigen konformen Abbildungen." *Acta Szeged*, **1**, 55–64 (1922).

ABSOLUTE SCALAR INVARIANTS AND THE ISOMETRIC CORRESPONDENCE OF RIEMANN SPACES

BY T. Y. THOMAS

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY

Communicated August 13, 1945

Let R_n be a Riemann space (defined by a positive definite quadratic differential form). Denote by $I_1, \ldots, I_p$ with $1 \leqq p \leqq n$ a set of absolute scalar differential invariants of R_n. These invariants determine absolute scalar functions $I_1(x), \ldots, I_p(x)$ of the coordinates $x^1, \ldots, x^n$ of R_n with functional matrix

$$\left\| \begin{matrix} \dfrac{\partial I_1}{\partial x^1} & \cdots & \dfrac{\partial I_1}{\partial x^n} \\ \cdots\cdots\cdots & & \\ \cdots\cdots\cdots & & \\ \dfrac{\partial I_p}{\partial x^1} & \cdots & \dfrac{\partial I_p}{\partial x^n} \end{matrix} \right\|.$$

If this matrix has rank p at all points of R_n the scalars $I_1, \ldots, I_p$ are functionally independent, and if the functional matrix of every $q > p$ scalar invariants has rank not exceeding p the above scalars $I_1, \ldots, I_p$ (whose functional matrix does not vanish over R_n) are said to constitute a fundamental set of scalar invariants of R_n. In the following we deal with conditions for the isometric correspondence of two Riemann spaces on the basis of their absolute scalar invariants. The discussion is of a local nature so that the term Riemann space as here used is identical with the concept of the neighborhood of a point of a Riemann space when the *im Grossen* viewpoint is adopted. The method depends primarily on the functional independence of scalar invariants rather than on the order of the differential invariants from which these scalar functions are derived. Differentiability requirements on the components $g_{\alpha\beta}(x)$ of the fundamental tensor of R_n are immediately obvious from the differential invariants employed in any instance and will not be stated explicitly.

A space R_n having a fundamental set of invariants $I_1, \ldots, I_p$ with $p \geqq 1$ will be said to be of category p. Obviously p cannot exceed n. In particular if all scalar differential invariants of R_n reduce to constants as functions of the coordinates we said that R_n is of category 0.

If R_n is of category p, where $0 \leqq p \leqq n$, and is in isometric correspondence with an $\bar{R}_n$, then $\bar{R}_n$ is of category p. For if $p = 0$, R_n is a space of constant curvature. Hence $\bar{R}_n$ must be of constant curvature and from this fact it follows readily that all scalar functions $\bar{I}(\bar{x})$ of $\bar{R}_n$ are constants. Hence $\bar{R}_n$ is of category 0. In general the $p(\geqq 1)$ independent scalars $I_1, \ldots, I_p$ of R_n will be carried into p independent scalars $\bar{I}_1, \ldots, \bar{I}_p$ of $\bar{R}_n$ by the correspondence; hence $\bar{R}_n$ cannot be of category $q < p$. Similarly if there were $q > p$ independent scalars in $\bar{R}_n$ these would yield $q > p$ independent scalars in R_n contrary to the hypothesis that R_n is of category p.

Let $I_1, \ldots, I_n$ be a fundamental set of scalar invariants of a space R_n of category n. For k fixed, the quantities $I_{k\alpha}$ (partial differentiation with respect to x^α) will be the components of a vector invariant of the space and hence the quantities $I_{jk} = g^{\alpha\beta} I_{j\alpha} I_{k\beta}$ will be absolute scalar invariants. Denoting by $\bar{I}_k$ and $\bar{I}_{jk}$ the corresponding scalars for a space $\bar{R}_n$ a necessary condition for the isometric correspondence of R_n and $\bar{R}_n$ is the consistency of the scalar equations

$$\bar{I}_j(\bar{x}) = I_j(x), \; \bar{I}_{jk}(\bar{x}) = I_{jk}(x) \tag{1}$$

where $j, k = 1, \ldots, n$. Let $x^\alpha = x^\alpha(\bar{x}^1, \ldots, \bar{x}^n)$ be a solution of (1). Then the $x^\alpha(\bar{x})$ will be differentiable functions; this follows from the theorem on the solution of implicit equations and the fact that the determinant $|I_{k\alpha}| \neq 0$ over R_n. Now from the first set of equations (1) we have

$$\bar{I}_{j\alpha}(\bar{x}) = I_{j\beta}(x) \frac{\partial x^\beta}{\partial \bar{x}^\alpha}.$$

From these relations and the second set of equations (1) we find

$$I_{j\mu} I_{k\nu} \left[\bar{g}^{\alpha\beta} \frac{\partial x^\mu}{\partial \bar{x}^\alpha} \frac{\partial x^\nu}{\partial \bar{x}^\beta} - g^{\mu\nu} \right] = 0.$$

Since $|I_{k\alpha}| \neq 0$ it follows that the bracket expression in the above equations must vanish. The solution $x^\alpha = x^\alpha(\bar{x})$ of (1) thus yields a (non-singular) isometric correspondence between R_n and $\bar{R}_n$. *Hence a space R_n of category n having a fundamental set of scalar invariants $I_1, \ldots, I_n$ can be put into isometric correspondence with a space $\bar{R}_n$ if, and only if, the equations (1) relating the coordinates x^α and $\bar{x}^\alpha$ of these spaces, admit a solution $x^\alpha = x^\alpha(\bar{x})$. If this solution exists it automatically gives the isometric correspondence.*

In what follows we limit our attention to spaces R_2. If R_2 is of category 2 the above result shows that five absolute scalars of R_2 suffice for the solution of the problem of the isometric correspondence between R_2 and any other space $\bar{R}_2$. Now suppose that R_2 is of category 1 and let J be the single independent scalar invariant of R_2, i.e., the matrix $\|\partial J/\partial x^\alpha\|$ has

rank 1 over R_2 (in particular J may be the Gaussian curvature K of R_2). Then form the scalars $J_1 = g^{\alpha\beta}J_\alpha J_\beta$ and $J_2 = g^{\alpha\beta}J_{\alpha\beta}$ where the $J_{\alpha\beta}$ are the components of the second covariant derivative of J.

LEMMA. *A space R_2 of category 1 having the fundamental scalar invariant J admits coordinates x, y (covering any point P of R_2) for which $J = x$ and the line-element has the form*

$$ds^2 = p(x)dx^2 + q(x)dy^2. \tag{2}$$

We can suppose $\partial J/\partial x^1 \neq 0$ at any point P of R_2 where x^1, x^2 are the coordinates of a suitably chosen system. Then $\eta^1 = J(x^1, x^2)$, $\eta^2 = x^2$ defines a non-singular transformation of a neighborhood of P. Let $g_{\alpha\beta} \rightarrow \theta_{\alpha\beta}(\eta)$ by this transformation. Now consider the equation

$$\theta^{1\beta}\frac{\partial v}{\partial \eta^\beta} = \theta^{11}\frac{\partial v}{\partial \eta^1} + \theta^{12}\frac{\partial v}{\partial \eta^2} = 0. \tag{3}$$

Since $\theta^{11} \neq 0$ this equation admits a solution $v(\eta^1, \eta^2)$ defined in the neighborhood of P, such that $\partial v/\partial \eta^2 \neq 0$ (see the discussion in C. Caratheodory, *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*, 1935, p. 24). Hence the transformation $u = \eta^1$, $v = v(\eta^1, \eta^2)$ is non-singular and if $\theta_{\alpha\beta} \rightarrow k_{\alpha\beta}(u, v)$ relative to the u, v coordinates, then $k^{12} = 0$ by (3) where the indices 1 and 2 refer to the variables u and v, respectively. Hence $k_{12} = 0$. Also $J = u$ and hence $J_1 = k^{11} = f(u)$ since J_1 is functionally dependent on J by hypothesis. Since $k^{11} = 1/k_{11}$ it follows that $k_{11} = p(u)$. We now have

$$J_2 = -J_\sigma\Gamma^\sigma_{\alpha\beta}k^{\alpha\beta} = -k^{\alpha\beta}\Gamma^1_{\alpha\beta} = -k^{11}\Gamma^1_{11} - k^{22}\Gamma^1_{22}, \tag{4}$$

where we have used the Γ's to denote the Christoffel symbols relative to the u, v system.

Now $k^{\alpha\beta}{}_{,\gamma} = 0$. Taking $\alpha = \beta = \gamma = 1$ the resulting equation shows that Γ^1_{11} is a function of u alone. Since J_2 is a function of u alone, it now follows that the quantity $k^{22}\Gamma^1_{22}$ in the right member of (4) is a function of u alone, i.e.,

$$k^{22}\Gamma^1_{22} = -\frac{1}{2}\phi(u), \tag{5}$$

or,

$$\frac{\partial \log k_{22}}{\partial u} = k_{11}\phi(u) = p(u)\phi(u) = \psi(u),$$

when we substitute the expression defining the Christoffel symbol Γ^1_{22} and replace the $k^{\alpha\beta}$ by their values in terms of the $k_{\alpha\beta}$ in (5). Integrating the

latter equation we see that h_{22} has the form $k_{22} = q(u)r(v)$. Then making the transformation

$$x = u, \; y = \int \sqrt{r(v)}\, dv$$

we obtain the form of the line-element given by (2), and $J = x$. This completes the proof of the lemma.

Now consider another space $\bar{R}_2$ of category 1. A necessary condition for the isometric correspondence of R_2 and $\bar{R}_2$ is that the three equations $\bar{J} = J$, $\bar{J}_1 = J_1$ and $\bar{J}_2 = J_2$ admit a solution $\bar{x}^\alpha = \bar{x}^\alpha(x^1, x^2)$ where $\bar{J}$, $\bar{J}_1$ and $\bar{J}_2$ denote the scalars in $\bar{R}_2$ which correspond to the scalars J, J_1 and J_2 in R_2. We assume the existence of this solution. Since $J_1 \neq 0$ in R_2 it follows that $\bar{J}_1 \neq 0$ in $\bar{R}_2$. Hence $\bar{J}$ is a fundamental scalar invariant for $\bar{R}_2$. Now suppose $P \rightarrow \bar{P}$ by the correspondence $\bar{x}^\alpha = \bar{x}^\alpha(x^1, x^2)$. Introduce the coordinates x, y in the neighborhood of the point P as given by the above lemma for the space R_2 and the corresponding coördinates $\bar{x}$, $\bar{y}$ in the neighborhood of $\bar{P}$ for the space $\bar{R}_2$ relative to which we have

$$\bar{J} = \bar{x}, \; d\bar{s}^2 = \bar{p}(\bar{x})d\bar{x}^2 + \bar{q}(\bar{x})d\bar{y}^2.$$

The above solution $\bar{x}^\alpha = \bar{x}^\alpha(x^1, x^2)$ implies the existence of a solution $\bar{x} = \bar{x}(x, y)$, $\bar{y} = \bar{y}(x, y)$ of the equations $\bar{J} = J$, $\bar{J}_1 = J_1$ and $\bar{J}_2 = J_2$ relative to these new coördinates. But from $\bar{J} = J$ we have $\bar{x} = x$. Also the relation $\bar{J}_1 = J_1$ yields $\bar{g}^{11}(\bar{x}) = g^{11}(x)$ from which we deduce $\bar{g}_{11}(\bar{x}) = g_{11}(x)$ or $\bar{p}(\bar{x}) = p(x)$. Then from $\bar{J}_2 = J_2$ we find

$$\bar{g}^{11}\bar{\Gamma}^1_{11} + \bar{g}^{22}\bar{\Gamma}^1_{22} = g^{11}\Gamma^1_{11} + g^{22}\Gamma^1_{22}. \tag{6}$$

Since g^{11} depends on x alone it follows that Γ^1_{11} depends on x alone. Similarly for $\bar{\Gamma}^1_{11}$. Also from the relations $\bar{g}^{11} = g^{11}$, $\bar{g}_{11} = g_{11}$ and $\bar{x} = x$ we obtain $\bar{\Gamma}^1_{11}(x) = \Gamma^1_{11}(x)$. Hence (6) reduces to $\bar{g}^{22}\bar{\Gamma}^1_{22} = g^{22}\Gamma^1_{22}$. On simplifying, this latter relation becomes

$$\frac{1}{\bar{q}}\frac{d\bar{q}}{d\bar{x}} = \frac{1}{q}\frac{dq}{dx}.$$

Then using the fact that $\bar{x} = x$ and integrating we have $\bar{q} = C^2 q$ where C is a constant.

From the relations $\bar{p}(\bar{x}) = p(x)$, $\bar{q}(\bar{x}) = C^2 q(x)$ and $\bar{x} = x$ we see immediately that the transformation $x = \bar{x}$, $y = C\bar{y} + d$ where d is a suitable constant gives an isometric correspondence between the neighborhoods of P and $\bar{P}$. Understanding that the term Riemann space is identical with a suitably restricted neighborhood in accordance with the local viewpoint the result obtained can now be stated as follows. *A Riemann space R_2 of category 1 with fundamental scalar invariant J can be put in isometric correspondence with a Riemann space $\bar{R}_2$ if, and only if, R_2 is of category 1 and*

the three scalar equations $\bar{J} = J$, $\bar{J}_1 = J_1$ *and* $\bar{J}_2 = J_2$ *admit a solution* $\bar{x}^\alpha = \bar{x}^\alpha(x^1, x^2)$ *relating the coördinates* x^α *of* R_2 *to the coördinates* $\bar{x}^\alpha$ *of* $\bar{R}_2$.

Since a Riemann space R_2 of category 0 has its Gaussian curvature $K =$ const. we can immediately state the following result. *A Riemann space* R_2 *of category 0 can be put in isometric correspondence with a Riemann space* $\bar{R}_2$ if, and only if, $\bar{R}_2$ *is of category 0 and* $\bar{K} = K$ *where* K *and* $\bar{K}$ *are the Gaussian curvatures of* R_2 *and* $\bar{R}_2$, *respectively.* It has already been observed that the categories of two Riemann spaces in isometric correspondence must be the same. The remainder of the proof consists in showing that two Riemann spaces of the same constant curvature can be put in isometric correspondence and this is a well-known result.

We leave open the problem of the extension of these results to Riemann spaces R_n where $n > 2$. Another problem, worthy of serious study, concerns the investigation of the isometric correspondence of Riemann spaces in the large (for example, open and compact spaces) on the basis of their absolute scalar invariants.

BERNOULLI'S NUMBERS AND CERTAIN ARITHMETIC QUOTIENT FUNCTIONS

By H. S. Vandiver

Department of Pure Mathematics, University of Texas

Communicated July 6, 1945

In another paper[1] the writer established a congruence involving Bernoulli numbers, which can be extended, as we shall now show, to the generalized Bernoulli numbers of the first order. Such a number, $b_n(g, h)$, is defined as the number obtained by expanding $(gb + h)^n$ in full by the binomial theorem and substituting b_k for b^k where b_k is obtained by the use of the formula $(b + 1)^i = b_a$, $i > 1$, and the left-hand member is expanded as before and b_k is substituted for b^k. Let p be an odd prime. Then $p - 1$ has the form mc, with c even. Now consider the expression

$$\sum_\rho (1 - (d\rho + 1)^{p-1})^2 \equiv N \pmod{p^2}, \qquad (1)$$

where ρ ranges over all the values such that $\rho^c \equiv 1 \pmod{p^2}$. Any term in this expression reduces to unity, modulo p^2, if ρ is such that

$$d\rho + 1 \equiv 0 \pmod{p}. \qquad (2)$$

For each ρ that does not satisfy (2), the corresponding term of (1) is divisible by p^2. In the relation $x^c \equiv 1 \pmod{p^2}$, the c incongruent solutions modulo p^2 are also incongruent modulo p. Hence, N in (1) is the number

of incongruent solutions ρ, modulo p, in (2) and hence equals 1 or 0, according as $d^c \equiv 1 \pmod{p}$ or $d^c \not\equiv 1 \pmod{p}$. Expansion of (1) gives

$$N \equiv c - 2c \sum_{k=0}^{m} d^{kc} \binom{p-1}{kc} + c \sum_{j=0}^{2m} d^{jc} \binom{2(p-1)}{jc}$$

modulo p^2, and this may be written

$$N \equiv -2c \sum_{k=1}^{m} d^{kc} \binom{p-1}{kc} + c \sum_{k=1}^{m} d^{kc} \binom{2(p-1)}{kc} + c \sum_{k=1}^{m} d^{(m+k)c} \binom{2(p-1)}{(m+k)c}.$$

Set $d = ga + h$ and multiply through by $(ga + h)^n$, with $n \not\equiv 0 \pmod{c}$. Using the value of N noted above, we have, after letting a range over the integers 0, 1, 2, ..., $p - 1$ and adding,

$$\sum_{r} (gr + h)^n \equiv -2c \sum_{k=1}^{m} S_{kc+n}(g, h) \binom{p-1}{kc} + c \sum_{k=1}^{m} S_{kc+n}(g, h) \binom{2(p-1)}{kc} + c \sum_{k=1}^{m} S_{mc+kc+n}(g, h) \binom{2(p-1)}{mc+kc} \quad (3)$$

modulo p^2, where

$$S_i(g, h) = \sum_{a=0}^{p-1} (ga + h)^i$$

and r ranges over the integers in the set 0, 1, 2..., $p - 1$ with $(gr + h)^c \equiv 1 \pmod{p}$. It can be shown, as follows, that for i even and $p > 3$

$$S_i(g, h) \equiv p b_i(g, h) \pmod{p^2}. \quad (4)$$

As noted in another[2] (relation (6), page 576 of the first reference) paper,

$$g(i + 1) S_i(g, h) = \sum_{q=0}^{i} \binom{i+1}{q+1} b_{i-q}(g, h) p^{q+1} g^{q+1}$$

which gives

$$S_i(g, h) = \sum_{q=0}^{i} \binom{i}{q} b_{i-q}(g, h) p \frac{p^{q-2}}{q+1} p^2 g^q. \quad (5)$$

For $q > 1$

$$p^{q-1} > (1 + 2)^{q-1} \geqq 1 + 2(q - 1) = 2q - 1 \geqq q + 1.$$

Hence each term of (5), beyond the second, is divisible by p^2, since the factor p cannot occur in the denominator of any fractional part oftener than in the numerator. But the second term will also be divisible by p^2, for i is

even and $b_{i-1}(g, h)$ is an integer.[2] This leaves (4). For i to be even in (4), n must be even, and the limitation $p > 3$ will be met; because if $p = 3$, then, as $p - 1 = cm$, $c = 2$, but $n \not\equiv 0 \pmod{c}$. Using (4) and dividing (3) through by p, we obtain, modulo p, with c and n even,

$$\frac{\sum_r (gr + h)^n}{p} \equiv -2c \sum_{k=1}^{m} \binom{p-1}{kc} b_{kc+n}(g, h) +$$
$$c \sum_{k=1}^{m} \binom{2(p-1)}{kc} b_{kc+n}(g, h) + c \sum_{k=1}^{m} \binom{2(p-1)}{mc + kc} b_{mc+kc+n}(g, h). \quad (6)$$

Reducing the binomial coefficients, modulo p, as on page 58 of a previous paper,[1] we obtain modulo p,

$$\frac{\sum_r (gr + h)^n}{p} \equiv -2c \sum_{k=1}^{m} (-1)^{kc} b_{kc+n}(g, h) +$$
$$c \sum_{k=1}^{m} (-1)^{kc}(kc + 1) b_{kc+n}(g, h) + c \sum_{k=1}^{m} (-1)^{kc-1} kc\, b_{p-1+kc+n}(g, h). \quad (7)$$

Now use the relation[3]

$$\frac{b_{s+p-1}(g, h)}{s + p - 1} \equiv \frac{b_s(g, h)}{s} \pmod{p}$$

where $g \not\equiv 0 \pmod{p}$ and for $s \not\equiv 0 \pmod{p-1}$, which latter limitation is here complied with since $n \not\equiv 0 \pmod{c}$. We may then write, modulo p, with $g \not\equiv 0 \pmod{p}$,

$$\frac{\sum_r (gr + h)^n}{p} \equiv c \sum_{k=1}^{m} (kc + 1) b_{kc+n}(g, h) -$$
$$c \sum_{k=1}^{m} kc \frac{kc + n - 1}{kc + n} b_{kc+n}(g, h) - 2c \sum_{k=1}^{m} b_{kc+n}(g, h) \equiv -cn \sum_{k=1}^{m} \frac{b_{kc+n}(g, h)}{kc + n}.$$

This gives

THEOREM I. *If p is an odd prime with $p - 1 = mc$, c even, and $g \not\equiv 0 \pmod{p}$, then*

$$\frac{\sum_r (gr + h)^n}{p} \equiv -cn \sum_{k=0}^{m-1} \frac{b_{kc+n}(g, h)}{kc + n} \pmod{p} \quad (8)$$

for n even and not divisible by c, where the summation on the left extends over all integers r in the set $0, 1, 2, \ldots, p - 1$, such that $(gr + h)^c \equiv 1 \pmod{p}$.

For the special case $g = 1$, $h = 0$ this is Theorem I of the previous[1] paper, and we shall now proceed to find another congruence involving $M(c, n, p)$, that symbol being defined as the left-hand member of (8) for this special

case. In another[4] paper the writer defined a repetitive set modulo l as a set of integers

$$r_1, r_2, \ldots, r_c \tag{9}$$

where l is any integer and $r_i < l$, such that there exists an $r \not\equiv 1$ with

$$rr_1, rr_2, \ldots, rr_c$$

being congruent in some order to (9) modulo l. Hence, for $l = p$, the integers r in the set $1, 2, \ldots, p-1$, which satisfy $r^c \equiv 1 \pmod{p}$, form a repetitive set modulo p with any such $r \not\equiv 1$ as multiplier. As the writer proved, the set of integers

$$\frac{y_a p + r_a}{r}; \quad a = 1, 2, \ldots, c$$

where

$$y_a \equiv -\frac{r_a}{p} \pmod{r}; \quad 0 \leqq y_a < r$$

is a permutation of (9). Hence

$$\sum_a r^n \left(\frac{y_a p + r_a}{r}\right)^n \equiv \sum_a r_a^n + np \sum_a y_a r_a^{n-1} \pmod{p^2}$$

or

$$(r^n - 1)M(c, n, p) \equiv n \sum_a y_a r_a^{n-1} \pmod{p}. \tag{10}$$

If

$$y_a \equiv -\frac{a}{p} \pmod{r}, \; 0 \leqq y_a < r,$$

then for the special case $c = p - 1$ (10) gives

$$n \sum_{a=1}^{p-1} y_a a^{n-1} \equiv (r^n - 1)\frac{S_n}{p} \pmod{p},$$

and it follows that

$$\sum_{a=1}^{p-1} y_a a^{n-1} \equiv \frac{(r^n - 1)b_n}{n} \pmod{p}, \tag{11}$$

for $n \not\equiv 0 \pmod{\overline{p-1}}$.

Let $n = cs + e$, $0 < e < c$, where s ranges over the values $0, 1, \ldots, m-1$, and let r belong to the exponent c modulo p. (11) gives on addition

$$(r^e - 1)\left(\sum_{s=0}^{m-1} \frac{b_{cs+e}}{cs+e}\right) \equiv \sum_s \sum_{a_1} y_{a_1} a_1^{cs+e-1} + \sum_s \sum_{a_2} y_{a_2} a_2^{cs+e-1} \pmod{p} \tag{12}$$

where a_1 ranges over the a's satisfying $a^c \equiv 1 \pmod{p}$ and a_2 over the a's satisfying $a^c \not\equiv 1 \pmod{p}$.

We have

$$\sum_s a_1^{cs+e-1} = a_1^{e-1}(1 + a_1^c + a_1^{2c} + \ldots + a_1^{(m-1)c}) \equiv a_1^{e-1} m \pmod{p},$$

and

$$\sum_s a_2^{cs+e-1} = a_2^{e-1}(1 + a_2^c + a_2^{2c} + \ldots + a_2^{(m-1)c}) \equiv$$

$$a_2^{e-1} \frac{a_2^{p-1} - 1}{a_2^c - 1} \equiv 0 \pmod{p}$$

since $a_2^c \not\equiv 1$. Hence, the right hand member of (12) reduces modulo p to

$$\sum_{a_1} m y_{a_1} a_1^{e-1}$$

Multiplication by $-ce$ gives

$$-ce(r^e - 1)\left(\sum_{s=0}^{m-1} \frac{b_{cs+e}}{cs+e}\right) \equiv \sum_{a_1} e y_{a_1} a_1^{e-1}. \qquad (13)$$

But we also have (10). Since r belongs to the exponent c modulo p, and $0 < e < c$, and

$$\frac{b_{s+p-1}}{s+p-1} \equiv \frac{b_s}{s} \pmod{p}$$

for $s \not\equiv 0 \pmod{p-1}$, comparison of (13) and (10) also yields the above noted special case of Theorem I for $g = 1$, $h = 0$.

It is not clear how the argument used in this second proof of Theorem I of the former[1] paper can be extended to give a proof of Theorem I of the present paper, since if we attempt this extension it seems necessary to employ a relation in another paper[4] (last separate congruence on page 123) and that relation was proved under a number of restrictions which would not apply in the statement of our general theorem.

[1] These PROCEEDINGS, **31**, 55–60 (1945), Theorem I. On page 55 of this paper on the second line above (4), insert the factor $(-1)^n$ in the right-hand member. The relation (4) is subject to the condition $n \not\equiv k\mu$. On page 56, second line, read "are at most only $(p-3)/2$" in lieu of "are only $p-1$." In the first line above relation (6), insert 0 after the second congruence sign. In the eighth line above relation (7), read $a < p$ in lieu of $a < (p-1)$. On page 59, fourth line above relation (14), read $\Sigma_r r^n$ in lieu of $\Sigma_r r$. In connection with the last formula on page 60, add the condition $(n(n+c), p) = 1$.

[2] These numbers were considered by the writer in *Duke Math. Jour.*, **8**, 575–584 (1941), and also in *Trans. Am. Math. Soc.*, **51**, 512–519 (1942).

[3] These PROCEEDINGS, **28**, 326 (1942), relation (8) for $j = 1$.

[4] *Bull. Am. Math. Soc.*, **46**, 122 (1940).

THE TOTAL DIFFERENTIAL EQUATION FOR THE EXPONENTIAL FUNCTION IN NON-COMMUTATIVE NORMED LINEAR RINGS

BY ARISTOTLE D. MICHAL

CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated June 30, 1945

A normed linear ring[1] is a normed linear space (say with real multipliers) over which there exists an associative (but not necessarily commutative) function AB (additive and continuous in each variable) with values in the space. We shall assume the existence of a unit element I and assume that the space is a complete normed linear space (Banach space). We need not normalize the modulus of AB and take it as the number 1, nor need we normalize the norm of the unit element I and take it as the number 1. Examples of normed linear rings are sets of square matrices, quaternions and more generally any (real) associative algebra topologized with a norm topology in a variety of possible ways. Important infinite dimensional examples abound in the theory of integral equations and functionals.

The exponential function e^A in a normed linear ring is defined by

$$e^A = I + A + \frac{A^2}{2!} + \ldots + \ldots.$$

Since $\|AB\| \leq m\|A\|\,\|B\|$, where m is the modulus of the bilinear function AB, it follows that the exponential function in a normed linear ring is an entire analytic function.[2] The object of this paper is to characterize the exponential function in a normed linear ring by a total differential system in Fréchet differentials. More specifically, this characterization is embodied in the following theorem.

THEOREM. *The non-linear total differential system in Fréchet differentials of functions $z(A)$ with arguments and values in a normed linear ring N*

$$\left.\begin{aligned} \delta z(A) &= \textstyle\int_0^1 z((1-\xi)A)\,\delta A\, z(\xi A)\,d\xi \\ z(0) &= I, \text{ the unit of } N, \end{aligned}\right\} \tag{1}$$

has a unique entire analytic solution given by the exponential function $z(A) = e^A$ in N.

Let

$$z(A) = \sum_{i=0} z_i(A) \quad (z_i(A), \text{ a homogeneous polynomial of degree } i)$$

be an entire analytic solution of (1). Clearly $z_0(A) = I$ and by calculation $z_1(A) = A$, $z_2(A) = A^2/2!$. In this and in what follows, a theorem by the

author[3] on term by term Fréchet differentiation of a power series in normed linear spaces is used. The following recurrence formula holds for $n \geq 2$ and for all A, $\delta A \in N$:

$$(n+1)P_{n+1}(A, A, \ldots, A, \delta A) = \int_0^1 \{\delta A z_n(\xi A) + z_n((1-\xi)A)\delta A + \sum_{\substack{i+j=n \\ i,j \geq 1}} z_i((1-\xi)A)\delta A z_j(\xi A)\} d\xi, \quad (2)$$

where $P_{n+1}(A_1, A_2, \ldots, A_{n+1})$ is the polar of the polynomial $z_{n+1}(A)$. The truth of the theorem follows after some calculation with the aid of the recurrence relation (2).

The ideas that led the author to the differential system (1) for the particular case in which the normed linear ring N is a matric ring are discussed in the author's paper, "Differential Equations in Fréchet Differentials Occurring in Integral Equations," these PROCEEDINGS, **31**, 252–258 (1945).

The characterization of other elementary functions in a normed linear ring N, such as $\sin A$ and $\cos A$, can also be effected by total differential systems in Fréchet differentials. Similar results can also be given for normed linear rings without a unit. One can, for example, study the properties of the function defined by the expansion

$$A + \frac{A^2}{2!} + \frac{A^3}{3!} + \ldots + \ldots.$$

It should be pointed out that the differential system (1) is not of the type for which Michal and Elconin have given existence and uniqueness theorems.[4]

[1] A search through the literature seems to reveal that the first systematic study of normed linear rings (non-Hilbert variety) was made by Martin, R. S., and Michal, A. D., in "Some Expansions in Vector Space," *Jour. Math. Pures et Appl.*, **13**, 69–91 (1934). This paper was presented before the Los Angeles 1932 national meeting of the American Mathematical Society. The author has made an extensive use of normed linear rings in his studies on general differential geometry. See Michal, A. D., "General Differential Geometries and Related Topics," *Bull. Amer. Math. Soc.*, **45**, 529–563 (1939), where earlier references are included. See also Michal, A. D., and Mewborn, A. B., "Abstract Flat Projective Differential Geometry," *Acta Mathematica*, **72**, 259–281 (1940). For problems in mathematical analysis, besides those treated in the Michal-Martin paper, the reader is referred to Michal, A. D., and Hyers, D. H., "Second Order Differential Equations with Two Point Boundary Conditions in General Analysis," *Amer. Jour. Math.*, **58**, 646–660 (1936); and to Michal, A. D., and Elconin, V., "Completely Integrable Differential Equations in Abstract Spaces," *Acta Mathematica*, **68**, 71–107 (1937). The last paper also considers briefly some questions on ideals in normed linear rings. In recent years many important contributions have been made to the theory of normed linear rings by many authors in Russia as well as in America. The Russian mathematician I. Gelfand is especially to be mentioned in this connection. See his paper in *Rec. Math.* (*Mat. Sbornik*), **51**, 3–24 (1941). For commutative normed linear rings, see Lorch, E. R., *Bull. Amer. Math. Soc.*, **50**, 447–463 (1944).

[2] Cf. Michal-Martin paper, loc. cit.

[3] Michal, A. D., *A Theorem on the Fréchet Differentials of Power Series in Normed Linear Spaces* (unpublished).

[4] Cf. Michal-Elconin paper, loc. cit.

UNIVERSAL RATIONAL FUNCTIONS

BY E. T. BELL

DEPARTMENT OF MATHEMATICS, CALIFORNIA INSTITUTE OF TECHNOLOGY

Communicated July 17, 1945

1. *Positive* in all that follows shall mean *greater than zero.* The coefficients in all rational functions considered are integers (actually all positive in the examples given).

According to a customary definition, a form F is universal if, for integer values (positive, zero or negative) of all the variables in F, F represents all integers, positive, zero or negative. Another type of universality, as in Waring's problem, requires that a form represent all positive or zero integers for positive or zero values of the variables. Thus the sum of four squares is universal in this sense, but is not universal if zero values of the variables are excluded. A further type restricts the universality to some proper subset C of all the integers, for example, cubes.

We shall call the rational function F/G, where F, G are forms with positive integer coefficients, and F/G is in its lowest terms, *universal*, if, for positive integer values of all the variables in F, G, F/G represents all positive integers. If F/G for positive integer values of the variables represents all integers in the subset C, F/G will be called *universal for C.* The universal rational functions considered here also represent all negative integers when negative integer values of the variables are admitted. These particular universal rational functions therefore represent all integers except zero for integer values of the variables different from zero. A point of interest, however, is that positive integer values of the variables suffice for the representation of all positive integers.

2. An example of a universal rational function in four variables x, y, z, w is

$$\frac{xz(x^{n-1}w^{n+1} + z^{n-1}y^{n+1})}{y^{n-1}x^{n+1} + w^{n-1}z^{n+1}},$$

where n is an arbitrary constant positive integer;

$$\frac{x^{m+1}y^{n+1}u^{s+1} + z^{m+1}w^{n+1}v^{s+1}}{x^{m-1}y^{n-1}u^{s-1} + z^{m-1}w^{n-1}v^{s-1}},$$

in six variables x, y, z, u, v, w, is universal for squares, where m, n, s are

arbitrary constant positive integers. In $2s$ variables $x_1, \ldots, x_s, y_1, \ldots, y_s$, with the m_i, n_i arbitrary constant positive integers and

$$F(x_1, \ldots, x_s, y_1, \ldots y_s) \equiv \prod_{i=1}^{s} x_i^{m_i+1} + \prod_{i=1}^{s} y_i^{n_i+1},$$

$$G(x_1, \ldots, x_s, y_1, \ldots, y_s) \equiv \prod_{i=1}^{s} x_i^{m_i-1} + \prod_{i=1}^{s} y_i^{n_i-1},$$

F/G is universal for squares. With $s > 1$ and

$$F \equiv x_s \prod_{i=1}^{s-1} x_i^{m_i+1} + y_s \prod_{i=1}^{s-1} y_i^{n_i+1},$$

$$G \equiv y_s \prod_{i=1}^{s-1} x_i^{m_i-1} + x_s \prod_{i=1}^{s-1} y_i^{n_i-1},$$

F/G is universal. For even integers,

$$\frac{x^2r^2 + y^2s^2 + z^2t^2 + 3w^2u^2}{xr + ys + zt}$$

is universal. For fourth powers, $(x^5 + u^2v^2w)/(x + w)$ is universal. Illustrative of an infinity of universal rational functions F/G in which each of F, G is a binomial, $x^2w^2(y^2 + zw)/(x^5 + w^5)$ is universal.

3. The foregoing examples suggest that universal rational functions are much more easily found or constructed than are universal forms. That this is so, is evident from the remark that an elementary monomial represents all positive integers for positive integer values of the variables: the monomial $x_1^{a_1} \ldots x_s^{a_s}$, in the s independent variables $x_1, \ldots, x_s$, is *elementary* if at least one of the positive integers $a_1, \ldots, a_s$ is 1. Without loss of generality, we shall assume $a_1 \leq a_2 \leq \ldots \leq a_s$; $(a_1, a_2, \ldots, a_s)$ is then the *index* of the monomial.

If F/G is universal, $F = nG$, where n is an arbitrary positive integer, has a solution for positive integer values of all the variables $y_1, \ldots, y_t$ in F, G. Suitable F, G are found by representing n as an elementary monomial, say $n = n_1^{a_1} \ldots n_s^{a_s}$, of index $(a_1, \ldots, a_s)$. For a given index, the total number of such representations is readily determined as a combinatorial function of the exponents in the prime decomposition of n. The representation of n in F/G is then of the type

$$y_i = M_i(n_1, \ldots, n_s), \; i = 1, \ldots, t,$$

where the M_i are monomials in $n_1, \ldots, n_s$.

Write $F/G \equiv R(y_1, \ldots, y_t)$. From what precedes, it follows that

$$R(z_1, \ldots, z_t)R(w_1, \ldots, w_t) = R(z_1w_1, \ldots, z_tw_t)$$

has an infinity of solutions in monomials $z_1, \ldots, z_t, w_1, \ldots, w_t$. For if

$n = n_1^{a_1} \ldots n_s^{a_s}$, $m = m_1^{a_1} \ldots m_s^{a_s}$ are representations of the arbitrary positive integers n, m in a monomial of index $(a_1, \ldots, a_s)$, their product is represented as $(n_1m_1)^{a_1} \ldots (n_sm_s)^{a_s}$ in a monomial of the same index. If $z_i = M_i(n_1, \ldots, n_s)$, $w_i = M_i(m_1, \ldots, m_s)$ represent n, m, respectively, in R,

$$z_iw_i = M(n_1m_1, \ldots, n_sm_s)$$

for the representation of nm in R.

It follows also that if $R(x_1, \ldots, x_t)$ is a universal rational function, and $R_1, \ldots, R_t$ are any universal rational functions in any variables, then $R(R_1, \ldots, R_t)$ is a universal rational function. For each variable denotes a positive integer, and hence each is representable in any universal rational function.

Appropriate F, G may be constructed indefinitely from the disjunction (logical product) and conjunction (logical sum) of simple multiplicative diophantine equations. If the totally distributed form of the disjunction or conjunction can be written as $F = MG$ (or as $G = MF$), where M is an elementary monomial, and M, G (or M, F) have no common factor, then F/G (or G/F) is universal. The solution of $F/G = n$ (or of $G/F = n$), where n is an arbitrary positive integer, is then obtained by expressing n in the form M. If $F = MG$ (or $G = MF$) was obtained by disjunction, the values of the variables are found by solving any one of the equations in the disjunction; if the equation was obtained by conjunction, the equations are solved simultaneously. Since in either case the solutions are the general solutions, given in a necessary and sufficient number of independent integer parameters, the corresponding M represent all positive integers. The first example in the preceding section was constructed from the disjunction of $xy^n = zw^n$, $yz^n = wx^n$; the fifth, from the conjunction of $xr = ys = zt = wu$. Both may be considerably generalized.